

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:27:33 ; Search time 34.9726 seconds
(without alignments)
2434.679 Million cell updates/sec

Title: US-09-920-954-1

Perfect score: 3101
Sequence: 1 MRKKVFLSVLSAAAILSTV.....EQAVNPVGVQXSLAIVN 639

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
1: /SID52/gcgdata/geneseq/genesep-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/genesep-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/genesep-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/genesep-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/genesep-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/genesep-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/genesep-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/genesep-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/genesep-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/genesep-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/genesep-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/genesep-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/genesep-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/genesep-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/genesep-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/genesep-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/genesep-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/genesep-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/genesep-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/genesep-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/genesep-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/genesep-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/genesep-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3030	97.7	639	20	AA17089
2	3022	97.5	640	20	AA17090
3	3021	97.4	640	20	AA17091
4	3017	97.3	639	20	AA17087
5	3007	97.0	640	20	AA17088
6	2723	87.8	641	20	AAW89547
7	2408	77.7	636	20	AAW89548
8	2155	69.5	434	23	AAW50080
9	2155	69.5	434	23	AAW50081
10	2082	67.1	434	23	AAW50085

11	2060.5	66.4	433	23	AAW50086	Bacillus sp alkali
12	1994	64.3	434	23	AAW50090	Bacillus sp KSM-KP
13	1952.5	63.0	433	23	AAW50084	Bacillus sp SD-521
14	1948.5	62.8	433	23	AAW50082	Bacillus sp D6-(FE
15	1941.5	62.6	433	23	AAW50083	Bacillus sp Y-(FER
16	1940.5	62.6	433	13	AAW62774	Alkali-protease Ya
17	1940.5	62.6	433	19	AAW62774	Modified Bacillus
18	1940.5	62.6	433	21	AAW65698	Bacillus sp. Lion
19	1940.5	62.6	433	21	AAW69207	Amino acid sequenc
20	1940.5	62.6	433	21	AAW64619	Bacillus lion y en
21	1644.5	53.0	345	19	AAW62230	Subtilase jpi70 fr
22	1644.5	53.0	345	20	AAW21654	Thermococcus prote
23	451.5	14.6	659	18	AAW24121	MO9856926 Seq ID 1
24	451.5	14.6	659	20	AAW24122	Pyrococcus furiosu
25	408	13.2	654	18	AAW24129	Pyrococcus furiosu
26	408	13.2	654	20	AAW24129	Pyrococcus furiosu
27	398	12.8	659	18	AAW24123	Protease. Synthetic
28	391	12.6	522	18	AAW24122	Hyperthermostable
29	391	12.6	522	18	AAW24122	Hyperthermostable
30	391	12.6	522	18	AAW24122	Hyperthermostable
31	377.5	12.2	545	22	AAW94838	T. yonsei subclit
32	358	11.5	1079	22	AAW8180	Transglutaminase r
33	340	11.0	520	18	AAW13667	Fragment of dhp a g
34	340	11.0	734	18	AAW13667	Streptomyces virid
35	340	11.0	823	18	AAW13667	Bacillus sp. Tyl45
36	307.5	9.9	903	17	AAW87005	Hyperthermostable
37	299.5	9.7	903	17	AAW87007	Protease. Pyrococ
38	299.5	9.7	1398	17	AAW87008	Pyrococcus furiosu
39	299.5	9.7	1398	18	AAW24124	MO9856926 Seq ID 6
40	299.5	9.7	1398	18	AAW24124	F. balearicum CP70
41	292	9.4	699	20	AAW08471	Hyperthermostable
42	288.5	9.3	237	17	AAW87009	RP-III residual pr
43	286	9.2	806	13	AAW27481	Peptide sequence.
44	283.5	9.1	188	17	AAW87011	Thermococcus prote
45	283.5	9.1	188	18	AAW24126	

ALIGNMENTS

```
RESULT 1
AA17089
ID AA17089 standard: Protein: 639 AA.
XX
AC AA17089:
XX
XX 21-JUL-1999 (first entry)
XX
DE Bacillus alkaline protease.
XX
XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme:
XX washing composition; oxidizing agent.
XX
XX Bacillus sp.
XX
XX OS
XX
XX WO9918218-A1.
XX
XX 15-APR-1999.
XX
XX PD
XX
XX PF 07-OCT-1998; 98WO-JP04528.
XX
XX PR 07-OCT-1997; 97JP-0274570.
XX
XX (KAOS ) KAO CORP.
XX Hitomi J, Kageyama Y, Kubota H, Nomura M, Okuda M;
XX Saeki K, Shikata S, Takaiwa M;
XX WPI: 1999-287736/24.
XX N-PSDB: AAX37277.
XX
XX Alkali protease from Bacillus used in washing powders
XX
```

PS Disclosure: Page 53-58; 71pp; Japanese.

XX The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidising agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidising agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease.

XX Sequence 639 AA:

Query Match 97.7%; Score 3030; DB 20; Length 639;
Best Local Similarity 93.4%; Pred. No. 4.2e-238;
Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```

OY 1 MRKKVFLSVLSAAALISTVALXNPASGAXRFDLDFKGIQTDTDXKXGFSKXOTGAAPL 60
DB 1 MRKKVFLSVLSAAALISTVALXNPASGAXRFDLDFKGIQTDTDXKXGFSKXOTGAAPL 60
OY 61 LVESENVKLKGLKKLETPANNKLIHQFNGPILEETKQLEXTGAKLIDYIPYAYI 120
DB 61 LVESENVKLKGLKKLETPANNKLIHQFNGPILEETKQLEXTGAKLIDYIPYAYI 120
OY 121 VEEGVYXKXXXIEHVESVEPLPYXIDPOLFTGASXLYAKALDTKXNKEVOLRG 180
DB 121 VEEGVYXKXXXIEHVESVEPLPYXIDPOLFTGASXLYAKALDTKXNKEVOLRG 180
OY 121 VEEGVYXKXXXIEHVESVEPLPYXIDPOLFTGASXLYAKALDTKXNKEVOLRG 180
DB 121 VEEGVYXKXXXIEHVESVEPLPYXIDPOLFTGASXLYAKALDTKXNKEVOLRG 180
OY 181 IEIXIAOXXXNDXYITAPKPEYKVMNDVARGIYKADVAOSSVGLYOGQIVAAADGLDT 240
DB 181 IEIXIAOXXXNDXYITAPKPEYKVMNDVARGIYKADVAOSSVGLYOGQIVAAADGLDT 240
OY 181 IEIXIAOXXXNDXYITAPKPEYKVMNDVARGIYKADVAOSSVGLYOGQIVAAADGLDT 240
DB 181 IEIXIAOXXXNDXYITAPKPEYKVMNDVARGIYKADVAOSSVGLYOGQIVAAADGLDT 240
OY 241 GRNDSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVFG 300
DB 241 GRNDSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVFG 300
OY 241 GRNDSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVFG 300
DB 241 GRNDSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVFG 300
OY 301 STNDSXGGLGGLPSNLQTLFSAQXASGARIHTNSMGAANVAGATTTDSRNVDTYRKNDMT 360
DB 301 STNDSXGGLGGLPSNLQTLFSAQXASGARIHTNSMGAANVAGATTTDSRNVDTYRKNDMT 360
OY 301 STNDSXGGLGGLPSNLQTLFSAQXASGARIHTNSMGAANVAGATTTDSRNVDTYRKNDMT 360
DB 301 STNDSXGGLGGLPSNLQTLFSAQXASGARIHTNSMGAANVAGATTTDSRNVDTYRKNDMT 360
OY 361 IIFPAENEXPNNGGTISAPGAKNAITVGATENLRPSFGSADININVAOFSSRGPTKDG 420
DB 361 IIFPAENEXPNNGGTISAPGAKNAITVGATENLRPSFGSADININVAOFSSRGPTKDG 420
OY 361 IIFPAENEXPNNGGTISAPGAKNAITVGATENLRPSFGSADININVAOFSSRGPTKDG 420
DB 361 IIFPAENEXPNNGGTISAPGAKNAITVGATENLRPSFGSADININVAOFSSRGPTKDG 420
OY 421 IKPDVAPGTXILSARSSLAPDSEFWANHDSKYAYMGGTSMATPIVAGNVAOQLREHFVN 480
DB 421 IKPDVAPGTXILSARSSLAPDSEFWANHDSKYAYMGGTSMATPIVAGNVAOQLREHFVN 480
OY 421 IKPDVAPGTXILSARSSLAPDSEFWANHDSKYAYMGGTSMATPIVAGNVAOQLREHFVN 480
DB 421 IKPDVAPGTXILSARSSLAPDSEFWANHDSKYAYMGGTSMATPIVAGNVAOQLREHFVN 480
OY 481 RGTTPPSLLKALLAGAADXGLGYNGNOGMRVTLDSLVNAYVNESSALSTSOKATY 540
DB 481 RGTTPPSLLKALLAGAADXGLGYNGNOGMRVTLDSLVNAYVNESSALSTSOKATY 540
OY 481 RGTTPPSLLKALLAGAADXGLGYNGNOGMRVTLDSLVNAYVNESSALSTSOKATY 540
DB 481 RGTTPPSLLKALLAGAADXGLGYNGNOGMRVTLDSLVNAYVNESSALSTSOKATY 540
OY 541 XFPATGKRLKLSIWSADPASTTASVTLVNDLDTITAPNGXYXVNGNFXKXXXNMVG 600
DB 541 XFPATGKRLKLSIWSADPASTTASVTLVNDLDTITAPNGXYXVNGNFXKXXXNMVG 600
OY 541 XFPATGKRLKLSIWSADPASTTASVTLVNDLDTITAPNGXYXVNGNFXKXXXNMVG 600
DB 541 XFPATGKRLKLSIWSADPASTTASVTLVNDLDTITAPNGXYXVNGNFXKXXXNMVG 600
OY 601 RNNVENVFINSPOSGTYTIEVOAYNPVGPQNFSLAIYN 639
DB 601 RNNVENVFINSPOSGTYTIEVOAYNPVGPQNFSLAIYN 639

```

RESULT 2
AA17090
AA17090 standard: Protein: 640 AA.

XX
XX
XX

DT 21-JUL-1999 (first entry)
XX
XX Bacillus alkaline protease.
XX
XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW washing composition; oxidising agent.

OS Bacillus sp.
XX
XX MO9918218-21.

PN 15-APR-1999.

PD 07-OCT-1998; 98WO-3P04528.

PR 07-OCT-1997; 97JP-0274570.

XX (KAOS) KAO CORP.

PI Hitomi J, Kageyama Y, Kubota H, Nomura M, Okuda M;

PI Saeki K, Shikata S, Takaiwa M;

XX WPI: 1999-287736/24.

DR N-PSDB: AAX37278.

XX Alkali protease from Bacillus used in washing powders

PS Disclosure: Page 58-63; 71pp; Japanese.

CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidising agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidising agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease.

XX Sequence 640 AA:

Query Match 97.5%; Score 3022; DB 20; Length 640;
Best Local Similarity 93.3%; Pred. No. 1.9e-237;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

```

OY 2 RKKVFLSVLSAAALISTVALXNPASGAXRFDLDFKGIQTDTDXKXGFSKXOTGAAPL 61
DB 3 RKKVFLSVLSAAALISTVALXNPASGAXRFDLDFKGIQTDTDXKXGFSKXOTGAAPL 62
OY 62 VESENVKLKGLKKLETPANNKLIHQFNGPILEETKQLEXTGAKLIDYIPYAYI 121
DB 62 VESENVKLKGLKKLETPANNKLIHQFNGPILEETKQLEXTGAKLIDYIPYAYI 122
OY 63 VESENVKLKGLKKLETPANNKLIHQFNGPILEETKQLEXTGAKLIDYIPYAYI 122
DB 63 VESENVKLKGLKKLETPANNKLIHQFNGPILEETKQLEXTGAKLIDYIPYAYI 122
OY 122 EYEGDVXKXXXIEHVESVEPLPYXIDPOLFTGASXLYAKALDTKXNKEVOLRG 181
DB 122 EYEGDVXKXXXIEHVESVEPLPYXIDPOLFTGASXLYAKALDTKXNKEVOLRG 182
OY 123 EYEGDVXKXXXIEHVESVEPLPYXIDPOLFTGASXLYAKALDTKXNKEVOLRG 182
DB 123 EYEGDVXKXXXIEHVESVEPLPYXIDPOLFTGASXLYAKALDTKXNKEVOLRG 182
OY 182 EXIAQXXXNDXYITAPKPEYKVMNDVARGIYKADVAOSSVGLYOGQIVAAADGLDTG 241
DB 182 EXIAQXXXNDXYITAPKPEYKVMNDVARGIYKADVAOSSVGLYOGQIVAAADGLDTG 242
OY 242 RNDSSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVFG 301
DB 242 RNDSSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVFG 302
OY 302 RNDSSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVFG 302
DB 302 RNDSSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVFG 302
OY 302 RNDSSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVFG 302
DB 302 RNDSSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVFG 302
OY 303 IMDSXGGLGGLPSNLQTLFSAQXASGARIHTNSMGAANVAGATTTDSRNVDTYRKNDMT 362
DB 303 IMDSXGGLGGLPSNLQTLFSAQXASGARIHTNSMGAANVAGATTTDSRNVDTYRKNDMT 362

```

```

QY 362 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHYAOFSSNGPTKDGRI 421
DB 363 LFAAGNEXPNGGTISAPGTAKNAITVGATELRPSFGSYADNINHYAOFSSNGPTKDGRI 422
QY 422 KPDVMAPGTIIISARSSILAPDSSFWANHDSKYAVMGTSAPPIYAGNVAOLREHFVNR 481
DB 423 KPDVMAPGTIIISARSSILAPDSSFWANHDSKYAVMGTSAPPIYAGNVAOLREHFVNR 482
QY 482 GTPKPSLLKAAALIGAADXGLGTPNGOGKRVTLDDKSLNVAAYVNESSXLSQKATYX 541
DB 483 GTPKPSLLKAAALIGAADXGLGTPNGOGKRVTLDDKSLNVAAYVNESSXLSQKATYX 542
QY 542 FTATAGKPLKISLWSDAPASTTASVTLVNDLDLVTAPNGTYXVGNDEFPXKXNMDCR 601
DB 543 FTATAGKPLKISLWSDAPASTTASVTLVNDLDLVTAPNGTYXVGNDEFPXKXNMDCR 602
QY 602 NNVENVFINKPOSGTYTIEVOAYNVPVGPQXFSIATVN 639
DB 603 NNVENVFINKPOSGTYTIEVOAYNVPVGPQXFSIATVN 640

RESULT 3
AAV17091 standard: Protein: 640 AA.
AC AAV17091:
DE 21-JUL-1999 (first entry)
DE Bacillus alkaline protease.
KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KM washing composition; oxidizing agent.
OS Bacillus sp.
PN WO9918218-A1.
PD 15-APR-1999.
PE 07-OCT-1998; 98WO-JP04528.
PR 07-OCT-1997; 97JP-0274570.
PA (KAOS ) KAO CORP.
PI Hitomi J, Kageyama Y, Kubota H, Nomura M, Okuda M,
PI Seeki K, Shikata S, Takaiwa M,
DR WPI: 1999-287736/24.
DR N-PSDB; AAK37279.
PS Alkali protease from Bacillus used in washing powders
PS Disclosure: Page 63-68; 71pp; Japanese.
XX The invention relates to alkaline proteases produced by strains of
XX Bacillus. The proteases ability to digest casein is not inhibited by
XX oleic acid and they have a high stability to oxidizing agents. The
XX alkaline protease of the invention has the following properties: (a) it
XX is active over the pH range 4-13 and has at least 80% of its optimum
XX activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
XX stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)
XX its ability to digest casein is not inhibited by oleic acid; (e) it has
XX molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
XX used as enzymes in washing compositions for use in automatic dishwashers
XX and for washing clothes. The stability to oxidizing agents allows the
XX enzyme to be an effective component of washing compositions including
XX bleaches. The present sequence represents an alkaline protease.
SQ Sequence 640 AA:
Query Match 97.4%; Score 3021; DB 20; Length 640;

```

```

Best Local Similarity 93.3%; Pred. No. 2,2e-237;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
QY 2 RKKVFLSVLSAAILSTYALKNPAGAKXKFDIDFGIOTTTDDXGFSKQXOTAAFL 61
DB 3 RKKVFLSVLSAAILSTYALKNPAGAKXKFDIDFGIOTTTDDXGFSKQXOTAAFL 62
QY 62 VESENVKLLKGLKKLETPANNKLIHQFNGLIETKQXLEXTGAKILIDYIPDYAVIV 121
DB 63 VESENVKLLKGLKKLETPANNKLIHQFNGLIETKQXLEXTGAKILIDYIPDYAVIV 122
QY 122 EYEDGVASXXXIXEHVESVEPLPYRYXIDPOLFTKGASXLVKAXALDTQXKKEVQLRGI 181
DB 123 EYEDGVASXXXIXEHVESVEPLPYRYXIDPOLFTKGASXLVKAXALDTQXKKEVQLRGI 182
QY 182 EYEDGVASXXXIXEHVESVEPLPYRYXIDPOLFTKGASXLVKAXALDTQXKKEVQLRGI 181
DB 183 EYEDGVASXXXIXEHVESVEPLPYRYXIDPOLFTKGASXLVKAXALDTQXKKEVQLRGI 182
QY 242 RNDSSMHEAFRGKITALYALGRTNANDNGHGFHVGSVLNGXTNKGMAPOANLVFQS 301
DB 243 RNDSSMHEAFRGKITALYALGRTNANDNGHGFHVGSVLNGXTNKGMAPOANLVFQS 302
QY 302 IMDSXGGLGSPNLOTLEFSQAXSAGARIHTNSGAAVNCATTTDSRNDVYVRKNDMTI 361
DB 303 IMDSXGGLGSPNLOTLEFSQAXSAGARIHTNSGAAVNCATTTDSRNDVYVRKNDMTI 362
QY 362 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHYAOFSSNGPTKDGRI 421
DB 363 LFAAGNEXPNGGTISAPGTAKNAITVGATELRPSFGSYADNINHYAOFSSNGPTKDGRI 422
QY 422 KPDVMAPGTIIISARSSILAPDSSFWANHDSKYAVMGTSAPPIYAGNVAOLREHFVNR 481
DB 423 KPDVMAPGTIIISARSSILAPDSSFWANHDSKYAVMGTSAPPIYAGNVAOLREHFVNR 482
QY 482 GTPKPSLLKAAALIGAADXGLGTPNGOGKRVTLDDKSLNVAAYVNESSXLSQKATYX 541
DB 483 GTPKPSLLKAAALIGAADXGLGTPNGOGKRVTLDDKSLNVAAYVNESSXLSQKATYX 542
QY 542 FTATAGKPLKISLWSDAPASTTASVTLVNDLDLVTAPNGTYXVGNDEFPXKXNMDCR 601
DB 543 FTATAGKPLKISLWSDAPASTTASVTLVNDLDLVTAPNGTYXVGNDEFPXKXNMDCR 602
QY 602 NNVENVFINKPOSGTYTIEVOAYNVPVGPQXFSIATVN 639
DB 603 NNVENVFINKPOSGTYTIEVOAYNVPVGPQXFSIATVN 640

RESULT 4
AAV17087 standard: Protein: 639 AA.
AC AAV17087:
DE 21-JUL-1999 (first entry)
DE An alkaline protease sequence from Bacillus species.
KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KM washing composition; oxidizing agent.
OS Bacillus sp.
PN WO9918218-A1.
PD 15-APR-1999.
PE 07-OCT-1998; 98WO-JP04528.

```

```

XX 07-OCT-1997: 97JP-0274570.
PR (KAOS ) KAO CORP.
XX
PA Hitomi J, Kageyama Y, Kubota H, Nomura M, Okuda M;
PI Saeki K, Shikata S, Takaiwa M;
XX
DR WPI: 1999-287736/24.
DR N-PSDB: AAX37277.
XX
PT Alkali protease from Bacillus used in washing powders
XX
PS Claim 3; Page 47-50; 71pp; Japanese.
XX
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease of the
CC invention.
XX
XX Sequence 639 AA;
SQ
Query Match 97.3%; Score 3017; DB 20; Length 639;
Best Local Similarity 100.0%; Pred. No. 4, 8e-237;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTITDXXGFSKXQGTGAAP 60
DB 1 MRKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTITDXXGFSKXQGTGAAP 60
QY 61 LVESNNVLAAGLKKKLETPVANKKLIHXONGPILEETKXOLETKAKIIDYIPDAYI 120
DB 61 LVESNNVLAAGLKKKLETPVANKKLIHXONGPILEETKXOLETKAKIIDYIPDAYI 120
QY 121 VEYEGDVYXXXXXIEHVESVEPYLPXYXIDPOLFTKGASLYVKAALDTKXKNEVQLRG 180
DB 121 VEYEGDVYXXXXXIEHVESVEPYLPXYXIDPOLFTKGASLYVKAALDTKXKNEVQLRG 180
QY 181 IEXIAQXXXXNDVXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIYVAADTGLDT 240
DB 181 IEXIAQXXXXNDVXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIYVAADTGLDT 240
QY 241 GRNDSMEARFKRTALXALGRNNANDTNGHGHVAGSVLGNKXTKKGAPOANLVEQ 300
DB 241 GRNDSMEARFKRTALXALGRNNANDTNGHGHVAGSVLGNKXTKKGAPOANLVEQ 300
QY 301 SIMDSXGGLGILPSNLTQTLFQOAXSAGARITNSGAAVNGAYTTDSNNVDYVRKNDMT 360
DB 301 SIMDSXGGLGILPSNLTQTLFQOAXSAGARITNSGAAVNGAYTTDSNNVDYVRKNDMT 360
QY 361 ILPAAGNEXPNGGTISAFGTAKMAITVGATENLRPSFGSYVDNINHVAQFSSRGFTKGR 420
DB 361 ILPAAGNEXPNGGTISAFGTAKMAITVGATENLRPSFGSYVDNINHVAQFSSRGFTKGR 420
QY 421 IKPDVMAAGTYILSRSSLPADSSFWANHDSKYAMGTSNATPTVAANVQLRHFVK 480
DB 421 IKPDVMAAGTYILSRSSLPADSSFWANHDSKYAMGTSNATPTVAANVQLRHFVK 480
QY 481 RGIITPKPBLKAAALAGADXLGIPNGNQGWRVTLDDKSLINVAIVNSSXLSSTOKATY 540
DB 481 RGIITPKPBLKAAALAGADXLGIPNGNQGWRVTLDDKSLINVAIVNSSXLSSTOKATY 540
QY 541 XFTATAGPLKISLVWSDAPASTASTVTLVNDLVLITAPNGTXYVGNDPXXPXXXXXNDWG 600

```

```

DB 541 XFTATAGPLKISLVWSDAPASTASTVTLVNDLVLITAPNGTXYVGNDPXXPXXXXXNDWG 600
QY 601 RNNVENWFINKPQSGTYTIEVOAYNVVPQPOXFEIAYN 639
DB 601 RNNVENWFINKPQSGTYTIEVOAYNVVPQPOXFEIAYN 639
XX
RESULT 5
AA17088
ID AA17088 standard; protein; 640 AA.
XX
AC AA17088;
XX
DT 21-JUL-1999 (first entry)
XX
DE An alkaline protease sequence from Bacillus species.
XX
KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW washing composition; oxidizing agent.
XX
OS Bacillus sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..640
FT N-PSDB: AAX37278.
FT
XX
PN W0918218-A1.
XX
PD 15-APR-1999.
XX
PE 07-OCT-1998; 98WO-JP04528.
XX
PR 07-OCT-1997; 97JP-0274570.
XX
XX (KAOS ) KAO CORP.
XX
PI Hitomi J, Kageyama Y, Kubota H, Nomura M, Okuda M;
PI Saeki K, Shikata S, Takaiwa M;
XX
DR WPI: 1999-287736/24.
DR N-PSDB: AAX37278.
XX
PT Alkali protease from Bacillus used in washing powders
XX
PS Claim 3; Page 50-53; 71pp; Japanese.
XX
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease of the
CC invention.
XX
XX Sequence 640 AA;
SQ
Query Match 97.0%; Score 3007; DB 20; Length 640;
Best Local Similarity 100.0%; Pred. No. 3, 1e-236;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 KKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTITDXXGFSKXQGTGAAPLV 62
DB 4 KKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTITDXXGFSKXQGTGAAPLV 63

```

```

OY 63 ESENVKLAKGLKKLETPVANNKLIHIXFNGPILEETKQXLEXTGAKILDYIPDAVIVE 122
DB 64 ESENVKLAKGLKKLETPVANNKLIHIXFNGPILEETKQXLEXTGAKILDYIPDAVIVE 123
OY 123 YEGDVYXXXXXIEHVESVEPEYLPXYXIDPOLFTKGASXLVAXALDTRKQNKKEVQJRGIE 182
DB 124 YEGDVYXXXXXIEHVESVEPEYLPXYXIDPOLFTKGASXLVAXALDTRKQNKKEVQJRGIE 183
OY 183 XIAOXXXXNDVXYITAKPEKYMNDVARGIVKADVAOSYGLXGOGIIVAVADPTGIDTR 242
DB 184 XIAOXXXXNDVXYITAKPEKYMNDVARGIVKADVAOSYGLXGOGIIVAVADPTGIDTR 243
OY 243 NDSSMHEAFRGKITALALGRTNNANDTNGHGTIVAGSVLNGXTNKGAPOANLVFQSI 302
DB 244 NDSSMHEAFRGKITALALGRTNNANDTNGHGTIVAGSVLNGXTNKGAPOANLVFQSI 303
OY 303 MDSXGGLGGLPSNLQTLFSGASAGARIHTNSWGAANVNGAYTTDSRVNDYVRKNDMTIL 362
DB 304 MDSXGGLGGLPSNLQTLFSGASAGARIHTNSWGAANVNGAYTTDSRVNDYVRKNDMTIL 363
OY 363 FAAGNEXPNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHVAOFSSRGPTKDGRIK 422
DB 364 FAAGNEXPNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHVAOFSSRGPTKDGRIK 423
OY 423 PDVNAFGTXILSARSLAPDSFMANHDSKYAYMGTSMATPIYAGVNAQURHEFPKNG 482
DB 424 PDVNAFGTXILSARSLAPDSFMANHDSKYAYMGTSMATPIYAGVNAQURHEFPKNG 483
OY 483 ITPKPSILKALINGADGGLGYPNGNGMGRTLDKSLNVAAYVNESSXLSTOKATYF 542
DB 484 ITPKPSILKALINGADGGLGYPNGNGMGRTLDKSLNVAAYVNESSXLSTOKATYF 543
OY 543 TATGKPLKISLVMSDAPASTASVTLVNDLDTYTAENGXYVGNDFPXXXXNMDGN 602
DB 544 TATGKPLKISLVMSDAPASTASVTLVNDLDTYTAENGXYVGNDFPXXXXNMDGN 603
OY 603 NVENFTXIPQSGTYTTEVOAVVPGPOXESLAVN 639
DB 604 NVENFTXIPQSGTYTTEVOAVVPGPOXESLAVN 640

RESULT 6
ID AAM89547 standard; Protein; 641 AA.
AC AAM89547;
DT 12-APR-1999 (first entry)
DE Bacillus jpl170 protease.
KW Protease; detergent; surfactant; leather processing; debittering;
XX flavour.
OS Bacillus sp.
XX
XX Key Location/Qualifiers
XX Peptide 1..33
XX Region /note= "signal peptide"
XX Protein /note= "34..208"
XX /note= "prepro region"
XX /note= "209..641"
XX /note= "mature protein"
XX
XX WO9856927-A2.
XX
XX 17-DEC-1998.
XX
XX 09-JUN-1998; 98MO-US12005.
XX
XX 12-JUN-1997; 97US-0873479.
XX
XX (NOVO ) NOVO NORDISK BIOTECH INC.

```

```

XX Christianson L, Sloma A;
PI WPI; 1999-080908/07.
XX N-PSDB; AAV82382.
DR Novel protease from Bacillus subtilis 1620 - useful in laundry and
XX dishwashing detergents and for leather processing
XX
XX Claim 7; Page 53-54; 77pp; English.
XX
XX This is the amino acid sequence of a novel protease of Bacillus sp.
XX jpl170 (NC18 12513), as deduced from the nucleotide sequence of an
XX isolated gene (see AAV82382). The entire protein, including of an
XX signal peptide and prepro region, has 77% identity to alkaline
XX protease Y (see AAM89548) from Bacillus.
XX The invention provides
XX vectors, recombinant host cells and methods for the recombinant
XX production of the protease. The protease is used in laundry and
XX dishwashing detergents, for institutional and industrial cleaning,
XX and for leather processing, as well as for debittering and
XX enhancing the degree of hydrolysis of protein hydrolysates, for
XX flavour development through hydrolysis of proteins, degradation of
XX undesired peptides and in enzymatic synthesis of peptides. It has
XX enhanced stability towards oxidation under alkaline conditions,
XX e.g. towards bleaching agents of the peroxy type. The invention
XX also provides mutant cells in which the protease activity is
XX diminished. Such cells can be used for the production of
XX heterologous recombinant proteins.
XX
XX Sequence 641 AA:
XX
XX Query Match 87.8%; Score 2723; DB 20; Length 641;
XX Best Local Similarity 82.9%; Pred. No. 4,1e-213;
XX Matches 532; Conservative 40; Mismatches 66; Indels 4; Gaps 2;
XX
OY 1 MKK---KKVFLSVLSAAAILSTVALNPSAGAKRFDLPFGKIQTTTDXKXFSKQOTGA 57
DB 1 MKRKSKEVFLSVLSVALLSVALLSPSTIGANNFELDFGIELTLLEKAAKQCKTGK 60
OY 58 AAFVSENVKLKXGLKLETPVANNKLIHIXFNGPILEETKQXLEXTGAKILDYIPDY 117
DB 61 ASFLVNSENVKIPKSIQKLEVPADNKLVIYQFDGPILTEFQQLTEGTAKILDYIPDY 120
OY 118 AYIVEEGDVAXXXXXXIEHVESVEPEYLPXYXIDPOLFTKGASXLVAKAXALDTRKQNKREVO 177
DB 121 AYIVEYDGDVAKVNAIHAHLESEVEPYLPXYKIDPOLFTKGASSELVETVALDKKORSKEVR 180
OY 178 LRGIXIXIAQXXXXNDVXYITAKPEKYMNDVARGIVKADVAOSYGLXGOGIIVAAVDG 237
DB 181 LRGLEQIAQVATNNDVLYVTPKPEYELVDVARGIVKADVAONNFGILGCGQIYAAVDG 240
OY 238 LDTGRNDSMHEAFRGKITALALGRTNNANDTNGHGTIVAGSVLNGXTNKGAPOANLV 297
DB 241 LDTGRNDSMHEAFRGKITALALGRTNNANDTNGHGTIVAGSVLGN-ATNKGAPOANLV 299
OY 298 VFQSTIMDSXGGLGGLPSNLQTLFSGASAGARIHTNSWGAANVNGAYTTDSRVNDYVRKN 357
DB 300 VFQSTIMDSXGGLGGLPSNLQTLFSGASAGARIHTNSWGAANVNGAYTTDSRVNDYVRKN 359
OY 358 DMTILFAAGNEXPNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHVAOFSSRGPTK 417
DB 360 DMTILFAAGNEXPNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHVAOFSSRGPTK 419
OY 418 DGRIRPDVMAFGTXILSARSLAPDSFMANHDSKYAYMGTSMATPIYAGVNAQURHEF 477
DB 420 DGRIRPDVMAFGTXILSARSLAPDSFMANHDSKYAYMGTSMATPIYAGVNAQURHEF 479
OY 478 VKNRGITPKPSILKALIAAGADGGLGYPNGNGMGRTYTLKSLNVAAYVNESSXLSTSQK 537
DB 480 VKNRGVTPKPSILKALIAAGADGGLGYPNGNGMGRTYTLKSLNVAAYVNESSXLSTSQK 539
OY 538 ATYXFTATAGKPLKISLVMSDAPASTASVTLVNDLDTYTAENGXYVGNDFPXXXXNMDGN 597
DB 539 ATYXFTATAGKPLKISLVMSDAPASTASVTLVNDLDTYTAENGXYVGNDFPXXXXNMDGN 597

```


CC (34 - 38*) compared to prior art alkaline proteases (31 and 23*). This
 CC sequence represents a fragment of the alkaline protease Kp43 from
 CC Bacillus sp strain KSM-Kp43 which is used to create the modified protease
 CC represented in AAM50090.

XX Sequence 434 AA:

Query Match 69.5%; Score 2155; DB 23; Length 434;
 Best Local Similarity 96.3%; Pred. No. 4.3e-167;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 206 NDVARGIVKADVAQSSYGLXGOGQIVAAVADTGLDTRNDSSMHEAFRGKITLALYALGRN 265
 DB 1 NDVARGIVKADVAQSSYGLXGOGQIVAAVADTGLDTRNDSSMHEAFRGKITLALYALGRN 60
 QY 266 NANTNGHGHVAGSVLGNKGNKGMAPQANLVFQSIMDSXGGLGGLPSNLQTLFSQAXS 325
 DB 61 NANTNGHGHVAGSVLGNKGNKGMAPQANLVFQSIMDSXGGLGGLPSNLQTLFSQAXS 120
 QY 326 AGARIHTNSGGAAVNGAYTTDSRNVDVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
 DB 121 AGARIHTNSGGAAVNGAYTTDSRNVDVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
 QY 386 TVGATENLRPSFGSYADNINHAOFSSRGPTKDGRIKPDVMAFGTILSARSLAPDSSF 445
 DB 181 TVGATENLRPSFGSYADNINHAOFSSRGPTKDGRIKPDVMAFGTILSARSLAPDSSF 240
 QY 446 WANHDSKYAAMGCTSMATPIVAGNVAQLREHFVNKRGITPKPSILKAALAGADGIGLY 505
 DB 241 WANHDSKYAAMGCTSMATPIVAGNVAQLREHFVNKRGITPKPSILKAALAGADGIGLY 300
 QY 506 PNGNQGGRVTLDKSLNVAAYVNESSXLTSTQKATYXFTATAGKPLKISLVMSDAPASTTA 565
 DB 301 PNGNQGGRVTLDKSLNVAAYVNESSXLTSTQKATYXFTATAGKPLKISLVMSDAPASTTA 360
 QY 566 SVTLVNDLDLVITAPNGTXYVGNDFXPPXXXNMDCRRNNEVFINKPQSGTYTIEVOAYN 625
 DB 361 SVTLVNDLDLVITAPNGTXYVGNDFXPPXXXNMDCRRNNEVFINKPQSGTYTIEVOAYN 420
 QY 626 VPVGPOXFSLAIVN 639
 DB 421 VPVGPOXFSLAIVN 434

DB 421 VPVGPOXFSLAIVN 434

RESULT 9
 ID AAM50081 standard; protein: 434 AA.

AC AAM50081;

DT 12-AUG-2002 (first entry)

DE Bacillus sp KSM-KP9860 alkaline protease protein fragment.

KW Alkaline protease; detergent; laundry; bleaching; dishwasher.

OS Bacillus sp.

PN EPI209233-A2.

PD 29-MAY-2002.

PE 22-NOV-2001; 2001EP-0127851.

PR 22-NOV-2000; 2000JP-0355166.

PR 12-APR-2001; 2001JP-0114048.

XX (KAO) KAO CORP.

PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Arai H, Sumitomo N;
 PI Okuda M, Seki K;

DR WPI: 2002-437518/47.

XX New modified alkaline proteases useful in detergent compositions -
 PT Claim 5; Page 12-13; 25pp; English.

CC This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency &
 CC (34 - 38*) compared to prior art alkaline proteases (31 and 23*). This
 CC sequence represents a fragment of the alkaline protease KP9860 from
 CC Bacillus sp strain KSM-KP9860 described in the method of the invention.

XX Sequence 434 AA:

Query Match 69.5%; Score 2155; DB 23; Length 434;
 Best Local Similarity 96.3%; Pred. No. 4.3e-167;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 206 NDVARGIVKADVAQSSYGLXGOGQIVAAVADTGLDTRNDSSMHEAFRGKITLALYALGRN 265
 DB 1 NDVARGIVKADVAQSSYGLXGOGQIVAAVADTGLDTRNDSSMHEAFRGKITLALYALGRN 60
 QY 266 NANTNGHGHVAGSVLGNKGNKGMAPQANLVFQSIMDSXGGLGGLPSNLQTLFSQAXS 325
 DB 61 NANTNGHGHVAGSVLGNKGNKGMAPQANLVFQSIMDSXGGLGGLPSNLQTLFSQAXS 120
 QY 326 AGARIHTNSGGAAVNGAYTTDSRNVDVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
 DB 121 AGARIHTNSGGAAVNGAYTTDSRNVDVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
 QY 386 TVGATENLRPSFGSYADNINHAOFSSRGPTKDGRIKPDVMAFGTILSARSLAPDSSF 445
 DB 181 TVGATENLRPSFGSYADNINHAOFSSRGPTKDGRIKPDVMAFGTILSARSLAPDSSF 240
 QY 446 WANHDSKYAAMGCTSMATPIVAGNVAQLREHFVNKRGITPKPSILKAALAGADGIGLY 505
 DB 241 WANHDSKYAAMGCTSMATPIVAGNVAQLREHFVNKRGITPKPSILKAALAGADGIGLY 300
 QY 506 PNGNQGGRVTLDKSLNVAAYVNESSXLTSTQKATYXFTATAGKPLKISLVMSDAPASTTA 565
 DB 301 PNGNQGGRVTLDKSLNVAAYVNESSXLTSTQKATYXFTATAGKPLKISLVMSDAPASTTA 360
 QY 566 SVTLVNDLDLVITAPNGTXYVGNDFXPPXXXNMDCRRNNEVFINKPQSGTYTIEVOAYN 625
 DB 361 SVTLVNDLDLVITAPNGTXYVGNDFXPPXXXNMDCRRNNEVFINKPQSGTYTIEVOAYN 420
 QY 626 VPVGPOXFSLAIVN 639
 DB 421 VPVGPOXFSLAIVN 434

RESULT 10

ID AAM50085 standard; protein: 434 AA.

AC AAM50085;

DT 12-AUG-2002 (first entry)

DE Bacillus sp alkaline protease protein A-1 fragment.

KW Alkaline protease; detergent; laundry; bleaching; dishwasher.

OS Bacillus sp.

PN EPI209233-A2.

PD 29-MAY-2002.

PE 22-NOV-2001; 2001EP-0127851.

PR 22-NOV-2000; 2000JP-0355166.

PR 12-APR-2001; 2001JP-0114048.

XX (KAOS) KAO CORP.
 XX PA Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 XX PI Okuda M, Saeki K;
 XX DR WPI: 2002-437518/47.
 XX
 XX New modified alkaline proteases useful in detergent compositions -
 XX PT
 XX PS Claim 5; Page 18-19; 25pp; English.
 XX
 XX This invention describes novel *Bacillus* sp. alkaline proteases useful in
 XX CC detergent compositions, especially in laundry, bleaching or automatic
 XX CC dishwasher detergents. The novel proteases have an increased detergency %
 XX CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 XX CC sequence represents a fragment of the alkaline protease A-1 from
 XX CC *Bacillus* sp NCIB12289 described in the method of the invention.
 XX
 XX Sequence 434 AA;

Query Match 67.1%; Score 2082; DB 23; Length 434;
 Best Local Similarity 91.5%; Pred. No. 3,8e-161;
 Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

QY 206 NDVARGIVADVAOSSYGLYGOGIYAVADTGLDGRNDSMHEAFRGKITALYALGRIN 265
 DB 1 NDVARGIVADVAOSSYGLYGOGIYAVADTGLDGRNDSMHEAFRGKITALYALGRIN 60
 QY 266 NANDTNGHGTTHVAGSVLNGXTNKGAPQANLVFQSIMDSXGGLGGLPSNLOTLEFSQAYS 325
 DB 61 NANDPNGHGTTHVAGSVLNGXTNKGAPQANLVFQSIMDSXGGLGGLPSNLOTLEFSQAYS 120
 QY 326 AGARITHTNSMGAAYNCATYTDSDRNDDYRKNDMTLLFAAGNEXPNRGITSAAGTAKNAI 385
 DB 121 AGARITHTNSMGAAYNCATYTDSDRNDDYRKNDMTLLFAAGNEXPNRGITSAAGTAKNAI 180
 QY 386 TVGATENTLRPSFGSYADNINHAOFSSRGPTKDGRIKPDVMAPGTYILSARSILAPDSSF 445
 DB 181 TVGATENTLRPSFGSYADNINHAOFSSRGPTKDGRIKPDVMAPGTYILSARSILAPDSSF 240
 QY 446 WANHDSKYAYMGTSMAPIVAGNVAQLREHFKVKNRGITPKPSLLKAAIAGAADVGLGY 505
 DB 241 WANHDSKYAYMGTSMAPIVAGNVAQLREHFKVKNRGITPKPSLLKAAIAGAADVGLGY 300
 QY 506 PNGNOGMRVTLDKSLNVAAYVNESSXLSOKATYXFTTACKPLKISLWSDAPASTA 565
 DB 301 PNGNOGMRVTLDKSLNVAAYVNESSXLSOKATYXFTTACKPLKISLWSDAPASTA 360
 QY 566 SVTLVNDLDTVITAPNGTYYVGNDEXXPYXXMMDGRRNNVENFIINXPOSGTYTIEVOAYN 625
 DB 361 SVTLVNDLDTVITAPNGTYYVGNDFAPYDNMMDGRRNNVENFIINXPOSGTYTIEVOAYN 420
 QY 626 VPVGPQXFSALIVN 639
 DB 421 VPVGPQXFSALIVN 434

RESULT 11
 AAM50086
 ID AAM50086 standard; protein; 433 AA.
 XX
 XX AAM50086;
 XX
 XX 12-AUG-2002 (first entry)
 XX
 XX *Bacillus* sp alkaline protease protein A-2 fragment.
 XX
 XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 XX *Bacillus* sp.
 XX
 XX EP1209233-A2.
 XX
 XX

XX 29-MAY-2002.
 PD
 XX 22-NOV-2001; 2001EP-0127851.
 XX
 XX 22-NOV-2000; 2000JP-0355166.
 XX
 XX 12-APR-2001; 2001JP-0114048.
 XX
 XX (KAOS) KAO CORP.
 XX PA Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 XX PI Okuda M, Saeki K;
 XX DR WPI: 2002-437518/47.
 XX
 XX New modified alkaline proteases useful in detergent compositions -
 XX PT
 XX PS Claim 5; Page 20-21; 25pp; English.
 XX
 XX This invention describes novel *Bacillus* sp. alkaline proteases useful in
 XX CC detergent compositions, especially in laundry, bleaching or automatic
 XX CC dishwasher detergents. The novel proteases have an increased detergency %
 XX CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 XX CC sequence represents a fragment of the alkaline protease A-2 from
 XX CC *Bacillus* sp NCIB12513 described in the method of the invention.
 XX
 XX Sequence 433 AA;

Query Match 66.4%; Score 2060.5; DB 23; Length 433;
 Best Local Similarity 91.7%; Pred. No. 2.1e-159;
 Matches 398; Conservative 12; Mismatches 23; Indels 1; Gaps 1;

QY 206 NDVARGIVADVAOSSYGLYGOGIYAVADTGLDGRNDSMHEAFRGKITALYALGRIN 265
 DB 1 NDVARGIVADVAOSSYGLYGOGIYAVADTGLDGRNDSMHEAFRGKITALYALGRIN 60
 QY 266 NANDTNGHGTTHVAGSVLNGXTNKGAPQANLVFQSIMDSXGGLGGLPSNLOTLEFSQAYS 325
 DB 61 NANDPNGHGTTHVAGSVLNGXTNKGAPQANLVFQSIMDSXGGLGGLPSNLOTLEFSQAYS 119
 QY 326 AGARITHTNSMGAAYNCATYTDSDRNDDYRKNDMTLLFAAGNEXPNRGITSAAGTAKNAI 385
 DB 120 AGARITHTNSMGAAYNCATYTDSDRNDDYRKNDMTLLFAAGNEXPNRGITSAAGTAKNAI 179
 QY 386 TVGATENTLRPSFGSYADNINHAOFSSRGPTKDGRIKPDVMAPGTYILSARSILAPDSSF 445
 DB 180 TVGATENTLRPSFGSYADNINHAOFSSRGPTKDGRIKPDVMAPGTYILSARSILAPDSSF 239
 QY 446 WANHDSKYAYMGTSMAPIVAGNVAQLREHFKVKNRGITPKPSLLKAAIAGAADVGLGY 505
 DB 240 WANHDSKYAYMGTSMAPIVAGNVAQLREHFKVKNRGITPKPSLLKAAIAGAADVGLGY 299
 QY 506 PNGNOGMRVTLDKSLNVAAYVNESSXLSOKATYXFTTACKPLKISLWSDAPASTA 565
 DB 300 PNGNOGMRVTLDKSLNVAAYVNESSXLSOKATYXFTTACKPLKISLWSDAPASTA 359
 QY 566 SVTLVNDLDTVITAPNGTYYVGNDEXXPYXXMMDGRRNNVENFIINXPOSGTYTIEVOAYN 625
 DB 360 SVTLVNDLDTVITAPNGTYYVGNDFAPYDNMMDGRRNNVENFIINXPOSGTYTIEVOAYN 419
 QY 626 VPVGPQXFSALIVN 639
 DB 420 VPVGPQXFSALIVN 433

RESULT 12
 AAM50090
 ID AAM50090 standard; protein; 434 AA.
 XX
 XX AAM50090;
 XX
 XX 12-AUG-2002 (first entry)
 XX
 XX
 XX
 XX

DB	421	VPVGPOTFSLATVN	434
<hr/>			
RESULT	13		
ID	AAM50084		
XX	AAM50084	standard; protein; 433 AA.	
AC	AAM50084;		
XX			
DT	12-AUG-2002	(first entry)	
DE	Bacillus sp SD-521 (FERM BP-11162)	alkaline protease protein fragment.	
XX			
KW	Alkaline protease; detergent; laundry; bleaching; dishwasher.		
XX			
OS	Bacillus sp.		
PN	EPI209233-A2.		
PD	29-MAY-2002.		
PF	22-NOV-2001; 2001EP-0127851.		
PR	22-NOV-2000; 2000JP-0355166.		
PR	12-APR-2001; 2001JP-0114048.		
PA	(KAOS) KAO CORP.		
PI	Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;		
PI	Okuda M, Saeki K;		
DR	WPI; 2002-437518/47.		
XX			
PT	New modified alkaline proteases useful in detergent compositions -		
PS	Claim 5; Page 16-18; 25pp; English.		
CC	This invention describes novel Bacillus sp. alkaline proteases useful in		
CC	detergent compositions, especially in laundry, bleaching or automatic		
CC	dishwasher detergents. The novel proteases have an increased detergency &		
CC	(34 - 38%) compared to prior art alkaline proteases (31 and 23%). This		
CC	sequence represents a fragment of the alkaline protease SD-521 from		
CC	Bacillus sp strain SD-521 (FERM BP-11162) described in the method of the		
CC	Invention.		
SQ	Sequence 433 AA;		
<hr/>			
Query Match	63.0%; Score 1952.5; DB 23; Length 433;		
Best Local Similarity	86.6%; Pred. No. 1.3e-150;		
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;			
OY	NDVAGIYKADVAQAQSSYGLYGOGQIVAAVDLTGIDGRNDSMHAEAFRGKITALVALGRTN	265	
DB	1 NDVAGIYKADVAQNNTYGLTGGQYVAADVTDLTDGRNDSMHAEAFRGKITALLALGRIN	60	
OY	NANDTINGHGTAVAGSVLGNGXTNKGMADPANLVFQSIMDSXGGLGPLSPNLQTLESQAXS	325	
DB	61 NANDPINGHTHVAGSVLGN-ALNKGMAPQANLVFQSIMDSXGGLGPLSPNLQLTSQAWN	119	
OY	AGARLHTMSWGCAAVNCATTTDSRRNDDVYRKDKMTILTPAAGEXENRGITISAPGAKNAI	385	
DB	120 AGARLHTMSWGCAVNVCAVTANSRDVEYVRNNDMTVLFPAAGNEGPSGITISAQTAKNAI	179	
OY	TVGATELENRPFSFGSYADIINIHNVAQFSSRGPTDGRIKIPDYMAPGTYIIARSLSLPDSSF	445	
DB	180 TVGATELENRPFSFGSLADNPNHIAQFSSSGKATRDGRIKIPDYVAPGFITIIARSLSLPDSSF	239	
OY	WANHDSKYAVMGCTSMATPIYAGNVAQOLREHFVKNRGITPKPSLLKAALIAGAADXCGLCY	505	
DB	240 WANYSKYAVMGCTSMATPIYAGNVAQOLREHIKIRGITPKPSLLKAALIACATDVGCIGY	299	
OY	PENGNGKGRVTLDKSLNVAYNNESSXLSTSOAKATYXFYTATACKPLKISLWSADAPASTTA	565	

```

Db      300  PEGDQGMGVITLDDKSLNVAIVYNEATLALATGQKATYSEFOAQAGPELKLISLWMDAPGSTTA 359
QY      566  SVTLVNDLDLVITAPNGTXYVGNDEFPXPPXXNMWGGNNVENFVFNXPQSGTYTIEVOAYN 625
      1  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      360  SYTLVNDLDLVITAPNGQXYVGNDEFSYPDNNMMDGNNVENFVFNAPQSGTYTIEVOAYN 419
QY      626  VPVYGPQXESLALVYN 639
      11  |||  |||||  :
Db      420  VPSGPQRFSLAIYH 433

RESULT 14
AAM50082
ID      AAM50082 standard; protein: 433 AA.
AC      AAM50082;
XX      AAM50082;
XX      12-AUG-2002 (first entry)
DT      12-AUG-2002 (first entry)
DE      Bacillus sp D6-(FERM P1592) alkaline protease protein fragment.
XX      Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX      Bacillus sp.
OS      Bacillus sp.
XX      EP1209233-A2.
XX      29-MAY-2002.
XX      22-NOV-2001; 2001EP-0127851.
PF      22-NOV-2000; 2000JP-0355166.
PR      12-APR-2001; 2001JP-0114048.
XX      (KAOS ) KAO CORP.
PA      Hataeda Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
XX      PI      Okuda M, Saeiki K;
XX      DR      WPI: 2002-437518/47.
XX      PT      New modified alkaline proteases useful in detergent compositions -
XX      Claim 5; Page 13-15; 25pp; English.
XX      PS      This invention describes novel Bacillus sp. alkaline proteases useful in
XX      CC      detergent compositions, especially in laundry, bleaching or automatic
XX      CC      dishwasher detergents. The novel proteases have an increased detergency &
XX      CC      (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX      CC      sequence represents a fragment of the alkaline protease E-1 from
XX      CC      Bacillus sp strain D6-(FERM-P1592) described in the method of the
XX      CC      invention.
XX      SQ      Sequence 433 AA:

Query Match      62.8%; Score 1948.5; DB 23; Length 433;
Best Local Similarity 86.4%; Pred. No. 2.8e-150;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1

QY      206  NDVARGYKADVAQSSYGLYGGGQIYAVADTGLDGRNDSSMHEAFRGTITLYALGRPN 265
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1  NDVARGIVADVAQNNYGLYGOGVAVAVADTGLDGRNDSSMHEAFRGTITLYALGRPN 60
QY      266  NANDPNGHETHAVGSYVLGNKGTNKGAPQANLVFOSIMDSXGGLGSLPNTLTPSOAXS 325
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      61  NANDPNGHETHAVGSYVLGN-ALNKGAPQANLVFOSIMDSXGGLGSLPNTLTPSOAWN 119
QY      326  AGARHTNSWGAAYVCAVYTDTSRNVDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      120  AGARHTNSWGAAPVCAVYANSROYDEYRNNDMYTLFAAGNEXPNGGTISAPGTAKNAI 179
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      386  TVGATENTLRPSFGSYADNINHVAQFSSRGPTRKGRIKPDVMAPGTXIISARSLAPDSSF 445
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

Db 180 TVGATENYRPSFGSIADNPNIHIOFSSRGATRDGRIRKPDVTAPGFTILSARSSLAPDSSF 239
Qy 446 WANHDSKYAYMGCTSMATPIYAGNVAOLREHFVNKRGITPRKPSLLKALILIAGADXLGY 505
Db 240 WANNSKYAYMGCTSMATPIYAGNVAOLREHFVNKRGITPRKPSLLKALILIAGADVLGY 299
Qy 506 PNGNOGMRVTLDKSLNVAAYVNESSXLSTSQKATYXFTATAGKPLKISLWMSDAPASTTA 565
Db 300 PNGDQGMGRVTLDKSLNVAAYVNEATALTGOKATYSPQTGACPKLISLWTDAPGSTTA 359
Qy 566 SYTLVNDLIDLVTAPNGCTXYVGNDFXXXNMDGRNNVENYFINXPOSGTYTIEVOAYN 625
Db 360 SYTLVNDLIDLVTAPNGOKYVGNDFSYPDNNMDGRNNVENYFINAPOSGTYTIEVOAYN 419
Qy 626 VPVGPQXFSLAIVN 639
Db 420 VPSGPQRFSLAIYH 433

RESULT 15
AAM50083
ID AAM50083 standard; protein: 433 AA.
XX
AC AAM50083;
XX
DF 12-AUG-2002 (first entry)
XX
DE Bacillus sp Y-(FERM BP-1029) alkaline protease protein fragment.
XX
KM Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX
OS Bacillus sp.
XX
PN EPI209233-A2.
XX
PD 29-MAY-2002.
XX
PF 22-NOV-2001; 2001EP-0127851.
XX
PR 22-NOV-2001; 2000JP-0355166.
PR 12-APR-2001; 2001JP-0114048.
XX
PA (KAOS) KAO CORP.
XX
PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
PI Okuda M, Saeiki K.
XX
DR WPI; 2002-437518/47.
XX
PT New modified alkaline proteases useful in detergent compositions -
XX
PS Claim 5; Page 15-16; 25pp; English.
XX
CC This invention describes novel Bacillus sp. alkaline proteases useful in
CC detergent compositions, especially in laundry, bleaching or automatic
CC dishwasher detergents. The novel proteases have an increased detergency &
CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
CC sequence represents a fragment of the alkaline protease Ya from
CC Bacillus sp strain Y-(FERM BP-1029) described in the method of the
CC invention.
XX
SQ Sequence 433 AA:

Query Match 62.6%; Score 1941.5; DB 23; Length 433;
Best Local Similarity 86.2%; Pred. No. 1e-149;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

Qy 206 NDVARGIVKADVAQSSYGLXGGOIVANVADGLDTGRNDSMHEAFRGKITVALYALGRTN 265
Db 1 NDVARGIVKADVAQSSYGLXGGOIVANVADGLDTGRNDSMHEAFRGKITVALYALGRTN 60
Qy 266 NANDTNGHGHVAGSVLGGXTNKGMAPQANLVFOSIMDSXGGLGGLPSNLQTLFSSQAXS 325
||:||||||||||||| ||||||||||||||| |||||||||||

Db 61 NASDPNGHGHVAGSVLGN-ALNKGMAPQANLVFOSIMDSXGGLGGLPSNLQTLFSSQAXN 119
Qy 326 AGARITHTNSMGAANVGAYTTDTSRNVDDYVRKNDMTLLFAAGNEXPNGGTTISAPGTAKNAI 385
Db 120 AGARITHTNSMGAANVGAYTTANSRQVDEYVRNNDMTVLLFAAGNEXPNGGTTISAPGTAKNAI 179
Qy 386 TVGATENLRPSFGSYADNIINHVAOFSSRGPTKGRIRKPDVMAPGTXILSARSSLAPDSSF 445
Db 180 TVGATENYRPSFGSIADNPNIHIOFSSRGATRDGRIRKPDVTAPGFTILSARSSLAPDSSF 239
Qy 446 WANHDSKYAYMGCTSMATPIYAGNVAOLREHFVNKRGITPRKPSLLKALILIAGADXLGY 505
Db 240 WANNSKYAYMGCTSMATPIYAGNVAOLREHFVNKRGITPRKPSLLKALILIAGADVLGY 299
Qy 506 PNGNOGMRVTLDKSLNVAAYVNESSXLSTSQKATYXFTATAGKPLKISLWMSDAPASTTA 565
Db 300 PNGDQGMGRVTLDKSLNVAAYVNEATALTGOKATYSPQTGACPKLISLWTDAPGSTTA 359
Qy 566 SYTLVNDLIDLVTAPNGCTXYVGNDFXXXNMDGRNNVENYFINXPOSGTYTIEVOAYN 625
Db 360 SYTLVNDLIDLVTAPNGOKYVGNDFSYPDNNMDGRNNVENYFINAPOSGTYTIEVOAYN 419
Qy 626 VPVGPQXFSLAIVN 639
Db 420 VPSGPQRFSLAIYH 433

Search completed: April 1, 2003, 12:38:17
Job time : 36.9726 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:37:09 ; Search time 13.9891 Seconds
(without alignments)
1343.997 Million cell updates/sec

Title: US-09-920-954-1

Perfect score: 3101
Sequence: 1 MRKKVFLSVLSAAALSTV.....EVOATNVPGVGRXSLAIVN 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:*

1: /cgn2_6/ptodata/1/1aa/5a.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5a.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6a.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6a.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6b.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/6c.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3030	97.7	639	4	US-09-509-814A-4
2	3022	97.5	640	4	US-09-509-814A-6
3	3021	97.4	640	4	US-09-509-814A-8
4	3017	97.3	639	4	US-09-509-814A-1
5	3007	97.0	640	4	US-09-509-814A-2
6	2723	87.8	641	2	US-08-873-479-42
7	2417.5	78.0	635	2	US-08-873-479-43
8	1940.5	62.6	433	4	US-09-104-623A-4
9	1940.5	62.6	433	4	US-09-019-532A-1
10	451.5	14.6	659	4	US-08-894-818B-1
11	451.5	14.6	659	4	US-09-445-472-12
12	408	13.2	654	4	US-08-894-818B-35
13	408	13.2	654	4	US-09-445-472-16
14	398	12.6	659	4	US-08-894-818B-5
15	391	12.6	412	4	US-09-445-472-1
16	391	12.6	522	4	US-08-894-818B-3
17	391	12.6	522	4	US-09-445-472-4
18	340	11.0	520	4	US-09-000-016-7
19	340	11.0	520	4	US-09-514-340-7
20	340	11.0	734	4	US-09-000-016-4
21	340	11.0	734	4	US-09-514-340-4
22	340	11.0	823	4	US-09-000-016-2
23	340	11.0	823	4	US-09-514-340-2
24	299.5	9.7	903	1	US-08-750-532-1
25	299.5	9.7	1398	4	US-08-750-532-9
26	299.5	9.7	1398	4	US-08-894-818B-8
27	299.5	9.7	1398	4	US-09-445-472-6

28	288.5	9.3	237	1	US-08-750-532-18	Sequence 18, Appl
29	282	9.1	418	2	US-08-873-479-44	Sequence 44, Appl
30	270.5	8.7	397	1	US-08-434-255-2	Sequence 2, Appl1
31	270.5	8.7	397	1	US-08-434-255-2	Sequence 4, Appl1
32	270.5	8.7	397	1	US-08-459-967-2	Sequence 2, Appl1
33	270.5	8.7	397	1	US-08-459-967-4	Sequence 4, Appl1
34	270.5	8.7	397	1	US-08-460-327-2	Sequence 2, Appl1
35	270.5	8.7	397	1	US-08-460-327-4	Sequence 4, Appl1
36	270.5	8.7	397	1	US-08-459-871-2	Sequence 2, Appl1
37	270.5	8.7	397	1	US-08-459-871-4	Sequence 4, Appl1
38	264	8.5	370	1	US-08-434-255-6	Sequence 6, Appl1
39	264	8.5	370	1	US-08-459-967-6	Sequence 6, Appl1
40	264	8.5	370	1	US-08-460-327-6	Sequence 6, Appl1
41	264	8.5	370	1	US-08-459-871-6	Sequence 6, Appl1
42	255	8.2	280	1	US-08-434-255-8	Sequence 8, Appl1
43	255	8.2	280	1	US-08-459-967-8	Sequence 8, Appl1
44	255	8.2	280	1	US-08-460-327-8	Sequence 8, Appl1
45	255	8.2	280	1	US-08-459-871-8	Sequence 8, Appl1

ALIGNMENTS

```
RESULT 1
US-09-509-814A-4
; Sequence 4, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKATA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKET, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUW
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509, 814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4:
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-4
Query Match 97.7%; Score 3030; DB 4; Length 639;
Best Local Similarity 93.4%; Pred. No. 1.9e-254;
Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
OY 1 MRKKVFLSVLSAAALSTVALXNPSAGXARXFDLDFKGIQTITDXGFSKXOTGAANF 60
DB 1 MRKKVFLSVLSAAALSTVALNPSAGDARTFDLDFKGIQTITDVSGFSKOROTGAANF 60
OY 1 LVESENKLLKGLKLETPVANNKLIHXFNCGPILBETKQXLEXTGAILDIYIDVAYI 120
DB 1 LVESENKLLKGLKLETPVANNKLIHXFNCGPILBETKQXLEXTGAILDIYIDVAYI 120
OY 121 VEYEGDVYKXXXXIEHVESVEPYLPXYIDPOLFTKGASXLVAXMLDTRKXNKRVOLRG 180
DB 121 VEYEGDVYKXXXXIEHVESVEPYLPXYIDPOLFTKGASXLVAXMLDTRKXNKRVOLRG 180
OY 181 IETXAOXXXXNDVXYTAKPEYKVMNDVARGIVADVAOSSYGLYGGOGIYVADTGLDT 240
DB 181 IETXAOXXXXNDVXYTAKPEYKVMNDVARGIVADVAOSSYGLYGGOGIYVADTGLDT 240
OY 241 GRNDSMHEAFROKITALYALGRTNANDTNGHGTIVAGSVLNGXTNKGAAPQANLVFO 300
```

```
|||||
Db 241 GNDSSMHEAFRGKITALTALGRTNANDTNGHGVAGSVLGCNGATNGMAPOANLVFQ 300
QY 301 STMDSSGGLGGLPSNLQTLFSSQAXSAGARIHTNSMGAANVATTDSSRVNDVYRRNDMT 360
Db 301 STMDSSGGLGGLPSNLQTLFSSQAXSAGARIHTNSMGAANVATTDSSRVNDVYRRNDMT 360
QY 361 LFAAGNEKPNNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAOFSSRGPKDGR 420
Db 361 LFAAGNEKPNNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAOFSSRGPKDGR 420
QY 421 IRPDVAPGTIILSARSSLAPDSSFANHDSKRAYMGTSMAPIYAGVAVOQLREHFVKN 480
Db 421 IRPDVAPGTIILSARSSLAPDSSFANHDSKRAYMGTSMAPIYAGVAVOQLREHFVKN 480
QY 481 RGTTPKPSLLKKAALIGAADXLGYPNGNGMGRTVLDKSLNAVYNESSXLSSTQKATY 540
Db 481 RGTTPKPSLLKKAALIGAADXLGYPNGNGMGRTVLDKSLNAVYNESSXLSSTQKATY 540
QY 541 XFTATGKPLKISLWSDAPASTTASVTLVNDLVTAPNGTXYVGNDFXXPKXNNMDG 600
Db 541 XFTATGKPLKISLWSDAPASTTASVTLVNDLVTAPNGTXYVGNDFXXPKXNNMDG 600
QY 601 RNNVENVFINXPOSGTYTTEVOAYNVPGPQXSLAIYN 639
Db 601 RNNVENVFINXPOSGTYTTEVOAYNVPGPQXSLAIYN 639
```

RESULT 2

```
US-09-509-814A-6
; Sequence 6, Application US/09509814A
; Patent No. 6376227
```

```
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-6
```

```
Query Match 97.5%; Score 3022; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 9.3e-254;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
```

```
QY 2 RKKKVFSLVSAAILSTVALNPSAGARXFDLDFKGIQTITDXXGFSKQXOTGAAPL 61
Db 3 KKKKVFSLVSAAILSTVALNPSAGARXFDLDFKGIQTITDXXGFSKQXOTGAAPL 62
QY 62 VESENVKLLKGLKKLETVPANNNKLIHIOFNGPILEETKQILEXTGAKILDIYIPYAYIV 121
Db 63 VESENVKLLKGLKKLETVPANNNKLIHIOFNGPILEETKQILEXTGAKILDIYIPYAYIV 122
QY 122 ESEGDVXSXXXIHEVESYEPYXIXIDPOLFTKGASLYVAXXALDTPQXKKEVOALRI 181
Db 123 ESEGDVKSATSTIEHVESYEPYXIXIDPOLFTKGASLYVAXXALDTPQXKKEVOALRI 182
```

```
QY 182 EXIAQXXXXNDYXITTAKEEYKVMNDVARGIVKADVAOSSYGLYGQGOIVAAVDTGLDTG 241
Db 183 EQIAQFAISNDVLYITTAKEEYKVMNDVARGIVKADVAOSSYGLYGQGOIVAAVDTGLDTG 242
QY 242 RNDSSMHEAFRGKITALTALGRTNANDTNGHGVAGSVLGCNGATNGMAPOANLVFQ 301
Db 243 RNDSSMHEAFRGKITALTALGRTNANDTNGHGVAGSVLGCNGATNGMAPOANLVFQ 302
QY 302 IMDSXGGLGGLPSNLQTLFSSQAXSAGARIHTNSMGAANVATTDSSRVNDVYRRNDMTI 361
Db 303 IMDSXGGLGGLPSNLQTLFSSQAXSAGARIHTNSMGAANVATTDSSRVNDVYRRNDMTI 362
QY 362 LFAAGNEKPNNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAOFSSRGPKDGR 421
Db 363 LFAAGNEKPNNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAOFSSRGPKDGR 422
QY 422 KPDVAPGTIILSARSSLAPDSSFANHDSKRAYMGTSMAPIYAGVAVOQLREHFVKN 481
Db 423 KPDVAPGTIILSARSSLAPDSSFANHDSKRAYMGTSMAPIYAGVAVOQLREHFVKN 482
QY 482 GTPKPSLLKKAALIGAADXLGYPNGNGMGRTVLDKSLNAVYNESSXLSSTQKATY 541
Db 483 GTPKPSLLKKAALIGAADXLGYPNGNGMGRTVLDKSLNAVYNESSXLSSTQKATY 542
QY 542 FTATGKPLKISLWSDAPASTTASVTLVNDLVTAPNGTXYVGNDFXXPKXNNMDG 601
Db 543 FTATGKPLKISLWSDAPASTTASVTLVNDLVTAPNGTXYVGNDFXXPKXNNMDG 602
QY 602 NNVENVFINXPOSGTYTTEVOAYNVPGPQXSLAIYN 639
Db 603 NNVENVFINXPOSGTYTTEVOAYNVPGPQXSLAIYN 640
```

RESULT 3

```
US-09-509-814A-8
; Sequence 8, Application US/09509814A
; Patent No. 6376227
```

```
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-8
```

```
Query Match 97.4%; Score 3021; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 1.1e-253;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
```

```
QY 2 RKKKVFSLVSAAILSTVALNPSAGARXFDLDFKGIQTITDXXGFSKQXOTGAAPL 61
Db 3 KKKKVFSLVSAAILSTVALNPSAGARXFDLDFKGIQTITDXXGFSKQXOTGAAPL 62
QY 62 VESENVKLLKGLKKLETVPANNNKLIHIOFNGPILEETKQILEXTGAKILDIYIPYAYIV 121
Db 63 VESENVKLLKGLKKLETVPANNNKLIHIOFNGPILEETKQILEXTGAKILDIYIPYAYIV 122
```


122	EXEGDVAXXXXXX	IVSVSEVPEYLPXXYXIDPOLFTGAGXILAKAALDTKXNKVEYLRGI	181	OTHER INFORMATION: xaa	is any amino acid
				NAME/KEY: misc_feature	
	LOCATION: (46)..(46)			OTHER INFORMATION: xaa	
123	EXEGDVKANTSTRIEIVSEVPEYLPTRIDPOLFTGASSELVKAVALDTKXNKVEYLRGI	182	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (47)..(47)			OTHER INFORMATION: xaa	
182	EXIAQXXXSNVYXITAKPEKXYMNDVARGTYKADVAASSYGLIGCOQIYAADVDTGLDYG	241	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (47)..(47)			OTHER INFORMATION: xaa	
183	EQIAQPAISNDVLIITAKPEKXYMNDVARGTYKADVAASSYGLIGCOQIYAADVDTGLDYG	242	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (47)..(47)			OTHER INFORMATION: xaa	
242	RNDSSMHEAFRCKITAYALGRTNNANDNGHGTIVAGSVLGNGXTNKGAPQANLVFQS	301	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (70)..(70)			OTHER INFORMATION: xaa	
243	RNDSSMHEAFRCKITAYALGRTNNANDNGHGTIVAGSVLGNGXTNKGAPQANLVFQS	302	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (70)..(70)			OTHER INFORMATION: xaa	
302	IMDSXGGLGGLPSNLQTLFSGAAXSAGARIHTNSMGAANVCAVYTTDSRNVDYRKNDMTI	361	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (74)..(74)			OTHER INFORMATION: xaa	
303	IMDSXGGLGGLPSNLQTLFSGAAXSAGARIHTNSMGAANVCAVYTTDSRNVDYRKNDMTI	362	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (74)..(74)			OTHER INFORMATION: xaa	
362	LFPAAGXENXNGGTISAGTAKNAITVGATENLRPFSGSYANINHHVQFSRGPPTDGRIT	421	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (89)..(89)			OTHER INFORMATION: xaa	
363	LFPAAGXENXNGGTISAGTAKNAITVGATENLRPFSGSYANINHHVQFSRGPPTDGRIT	422	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (89)..(89)			OTHER INFORMATION: xaa	
422	KRPVYAPGXTIISARSSSLAPDSSFANHDSKXYAVMGGSMTPTVYANVQVLEHEHYKRR	481	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (102)..(102)			OTHER INFORMATION: xaa	
423	KRPVYAPGXTIISARSSSLAPDSSFANHDSKXYAVMGGSMTPTVYANVQVLEHEHYKRR	482	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (102)..(102)			OTHER INFORMATION: xaa	
482	GITPKPSLLKAAALAGAADXGLGTPNGNOGKGRVTLDKSLNAVATVNESSXLSTSOAKATYX	541	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (105)..(105)			OTHER INFORMATION: xaa	
483	GITPKPSLLKAAALAGAADXGLGTPNGNOGKGRVTLDKSLNAVATVNESSXLSTSOAKATYX	542	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (105)..(105)			OTHER INFORMATION: xaa	
542	FTATACKPLKISLWSDAPASTASVTLVNDLDLVTAPNGTXYVGDFFXXPXXXMMDDR	601	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (131)..(131)			OTHER INFORMATION: xaa	
543	FTATACKPLKISLWSDAPASTASVTLVNDLDLVTAPNGTXYVGDFFXXPXXXMMDDR	602	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (131)..(131)			OTHER INFORMATION: xaa	
602	NNVENFIFXNPOSQTYTIEVOAYNVVPGQXFLAIYN	639	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (146)..(146)			OTHER INFORMATION: xaa	
603	NNVENFIFXNPOSQTYTIEVOAYNVVPGQXFLAIYN	640	OTHER INFORMATION: xaa	is any amino acid	
				NAME/KEY: misc_feature	
	LOCATION: (146)..(146)			OTHER INFORMATION: xaa	

```

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

```

Query Match 97.3%; Score 3017; DB 4; Length 639;
 Best Local Similarity 100.0%; Pred. No. 2.5e-253;
 Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MKKKVFLSVLSAAIILSTVALXNPSAGXARXFDLDFKGIQTTTXXGFSKXOTGAANF 60
DB 1 MKKKVFLSVLSAAIILSTVALXNPSAGXARXFDLDFKGIQTTTXXGFSKXOTGAANF 60
QY 61 LVESENVKLXGKLKKLETPANNKLIHXOFNGPILEETKOKLEXTGAKIIDIYDIYAI 120
DB 61 LVESENVKLXGKLKKLETPANNKLIHXOFNGPILEETKOKLEXTGAKIIDIYDIYAI 120
QY 121 VEYEGDVXSSXXIIIEHVESVEPYLPXYIDPOLFTKGASXLVKAYALDTKOKNKEVOLRG 180
DB 121 VEYEGDVXSSXXIIIEHVESVEPYLPXYIDPOLFTKGASXLVKAYALDTKOKNKEVOLRG 180
QY 181 IEXIQXXXSNDVXYITAKPEKVMNDVARGIVKADVAOSSGLIGOGQIYAVADTGLDT 240
DB 181 IEXIQXXXSNDVXYITAKPEKVMNDVARGIVKADVAOSSGLIGOGQIYAVADTGLDT 240
QY 241 GRNDSMHEAFRGKITITAYALGRTNANDTNGHGHVAGSVLNGCKTNGMAPOANLVFQ 300
DB 241 GRNDSMHEAFRGKITITAYALGRTNANDTNGHGHVAGSVLNGCKTNGMAPOANLVFQ 300
QY 301 SIMDSXGLIGLPSNLOTLFSQAXSAGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDMT 360

```

```

DB 301 SIMDSXGGLGCLPSNLOTLFSQAXSAGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDMT 360
QY 361 ILFAAGNEXPNGGTISAPGTAKNAITVGATEMLRPSFSGYADNINHVAQFSSRGPTKGR 420
DB 361 ILFAAGNEXPNGGTISAPGTAKNAITVGATEMLRPSFSGYADNINHVAQFSSRGPTKGR 420
QY 421 IKPDVAPGTXTILSARSSLAPDSSFWANHDSKYAYMGTSNATPIVAGNVAQLREHFYKN 480
DB 421 IKPDVAPGTXTILSARSSLAPDSSFWANHDSKYAYMGTSNATPIVAGNVAQLREHFYKN 480
QY 481 KGITPKPSILKALLAGAADXGLGPNQOGWRVTLKSLNVAAYVNSSXLSSTOKATY 540
DB 481 KGITPKPSILKALLAGAADXGLGPNQOGWRVTLKSLNVAAYVNSSXLSSTOKATY 540
QY 541 XFTATAGKPLKISLWSDAPASTASVTLVNDLVTAPNGCTXVYVGNDFXPPXXNNDG 600
DB 541 XFTATAGKPLKISLWSDAPASTASVTLVNDLVTAPNGCTXVYVGNDFXPPXXNNDG 600
QY 601 RNNVENVEFNKPSGTYTIEVQAVNPVGPQXFSLAIVN 639
DB 601 RNNVENVEFNKPSGTYTIEVQAVNPVGPQXFSLAIVN 639

```

```

RESULT 5
US-09-509-814A-2
Sequence 2, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid

```

NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (75)..(75)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (90)..(90)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (103)..(103)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (106)..(106)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (129)..(129)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (134)..(134)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (147)..(147)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (149)..(149)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (161)..(161)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (166)..(166)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (173)..(173)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (184)..(184)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (190)..(190)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (195)..(195)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (287)..(287)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (307)..(307)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (325)..(325)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (370)..(370)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature

LOCATION: (432)..(432)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (502)..(502)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (532)..(532)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: xaa
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: xaa
us-09-509-814a-2

Query Match 97.0%; Score 3007; DB 4; Length 640;
Best Local Similarity 100.0%; Pred. No. 1.9e-252;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 KKKVPLSVLSAAAIISVVALNPSGAXARFPDIDPFKIGITDTTDDXXGFSKXOTGMAAFV 62
Db 4 KKKVPLSVLSAAAIISVVALNPSGAXARFPDIDPFKIGITDTTDDXXGFSKXOTGMAAFV 63
Oy 63 ESENVKLXGLKKKLEIYVPAKKKLIHYQFNGPILLETGXKXLEXTGAKLIDYIPDYATVE 122
Db 64 ESENVKLXGLKKKLEIYVPAKKKLIHYQFNGPILLETGXKXLEXTGAKLIDYIPDYATVE 123
Oy 123 YEGDVXSSXXXIIEHVESVPEYLPYXXIDPOLFTKGASXLVKAAXALDTKQXKEVQLRGIE 182
Db 124 YEGDVXSSXXXIIEHVESVPEYLPYXXIDPOLFTKGASXLVKAAXALDTKQXKEVQLRGIE 183
Oy 183 XIAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGOGQIVAAVADTGLPTGR 242
Db 184 XIAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGOGQIVAAVADTGLPTGR 243
Oy 243 NDSSMEAFRGKITALVYALGRTNNANDTNGHTRHVGAVSYLGNXGXTNKGAPQANLVFQSI 302
Db 244 NDSSMEAFRGKITALVYALGRTNNANDTNGHTRHVGAVSYLGNXGXTNKGAPQANLVFQSI 303
Oy 303 MDSXGIGLGPENLQTLFQSOAXSAGARITHNSGCAVNCAYTTDSRNVDDVYRRKNDMTIL 362
Db 304 MDSXGIGLGPENLQTLFQSOAXSAGARITHNSGCAVNCAYTTDSRNVDDVYRRKNDMTIL 363
Oy 363 FAANGEXPNGGTTISAPGAKNAITVGAETEMLRPSFGSYADNTNHHVAQFSSNGPPTKGRIR 422
Db 364 FAANGEXPNGGTTISAPGAKNAITVGAETEMLRPSFGSYADNTNHHVAQFSSNGPPTKGRIR 423
Oy 423 PDVAPGTIILSARSSLAPDSSFANHDSRYAYMGSTSMATPTIVAGNVAQLREHVFVNRG 482
Db 424 PDVAPGTIILSARSSLAPDSSFANHDSRYAYMGSTSMATPTIVAGNVAQLREHVFVNRG 483

QY 483 ITPKPSLKAALAGAADXGLGYPNGNGWRVTLDKSLNVAVNESSXLSSTQKATYX 542
D 484 ITPKPSLKAALAGAADXGLGYPNGNGWRVTLDKSLNVAVNESSXLSSTQKATYX 543
QY 543 TATGKPLKISLWSDAPASTTASTVTLVNDLVTAPNGTXVYVNDXXXPKXXMMDCRN 602
D 544 TATGKPLKISLWSDAPASTTASTVTLVNDLVTAPNGTXVYVNDXXXPKXXMMDCRN 603
QY 603 NVENFINXPOSGTYTIEVQAVNVPVGPQXFSIAIVN 639
D 604 NVENFINXPOSGTYTIEVQAVNVPVGPQXFSIAIVN 640

RESULT 6

US-08-873-479-42
; Sequence 42, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873.479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agitis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-873-479-42

Query Match 87.8%; Score 2723; DB 2; Length 641;

Best Local Similarity 82.9%; Pred. No. 7.6e-228;

Matches 532; Conservative 40; Mismatches 66; Indels 4; Gaps 2;

QY 1 MRK--KKVFLSVLSAAALSTVALXNPSACXARXFDLDFKGIOTTTDXGFSKQXOTGA 57
D 1 MRKSGKRVFLSVLSVALLSVALSSPSTIGANNFELDFGIETLLEKAKATQKGTGK 60
QY 58 AAFVLESNNVLAAGLKKKLTTPANNKLIHXQFNGPILEETKQXLEXTGAKIIDYIPDY 117
D 61 ASFLVNSNNVLAAGLKKKLTTPANNKLIHXQFNGPILEETKQXLEXTGAKIIDYIPDY 120
QY 118 AYIYEYGDVYKXXIXLIEHVESVEPYLPXYIIDPOLFTKGSXLYKAAALDTQXNKREYQ 177
D 121 AYIYEYGDVYKXXIXLIEHVESVEPYLPXYIIDPOLFTKGSXLYKAAALDTQXNKREYQ 180

QY 178 LRGIEXIAQXXSNDVYXITAKPEYKVNVDVARGIVKADVAOSSYGLGQGOIYAAVADTG 237
D 181 LRGLQIAQVATNNDVLTVPKPEYEVLDVARGIVKADVAQNNFGLYGQGOIYAAVADTG 240
QY 238 LDTGNDSSMHEAFRGKITTALYALGRTNNANDTNHGHVAGSVLGNKXTKKGAPOANL 297
D 241 LDTGNDSSMHEAFRGKITTALYALGRTNNANDTNHGHVAGSVLGN-ATKKGMAPOANL 299
QY 298 VFOSIMDSXGIGLPSMLQTLFISOAXSAGARIHTNSGAENVNGAYTTDSRNVDDYVKN 357
D 300 VFOSIMDSXGIGLPSMLQTLFISOAXSAGARIHTNSGAENVNGAYTTDSRNVDDYVKN 359
QY 358 DMTILFAAGNEXPNGGTSAPGTAKNALTVGATENLRPSFGSYADNINHVAQFSSRGPTK 417
D 360 DMTILFAAGNEXPNGGTSAPGTAKNALTVGATENLRPSFGSYADNINHVAQFSSRGPTK 419
QY 418 DGRKIPDVMAFGTYLXLSRSSLAPDSSFWANHDSKYATMGSTSMATPIVAQNVQALRHF 477
D 420 DGRKIPDVMAFGTYLXLSRSSLAPDSSFWANHDSKYATMGSTSMATPIVAQNVQALRHF 479
QY 478 VKNRGITPKPSILKRALIAGAADXGLGYPNGNGWRVTLDKSLNVAVNESSXLSSTQK 537
D 480 VKNRGITPKPSILKRALIAGAADXGLGYPNGNGWRVTLDKSLNVAVNESSXLSSTQK 539
QY 538 ATYXFTATAGKPLKISLWSDAPASTTASTVTLVNDLVTAPNGTXVYVNDXXXPKXXN 597
D 540 ATYXFTATAGKPLKISLWSDAPASTTASTVTLVNDLVTAPNGTXVYVNDXXXPKXXN 599
QY 598 WDGRNNVENVFINXPOSGTYTIEVQAVNVPVGPQXFSIAIVN 639
D 600 WDGRNNVENVFINXPOSGTYTIEVQAVNVPVGPQXFSIAIVN 641

RESULT 7

US-08-873-479-43
; Sequence 43, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873.479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agitis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear


```

; EARLIER APPLICATION NUMBER: 0754/97
; EARLIER FILING DATE: 1997-06-25
; EARLIER APPLICATION NUMBER: 60/051,381
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: PCT/DK98/00015
; EARLIER FILING DATE: 1998-01-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
; US-09-019-532-4

```

```

Query Match          62.6% Score 1940.5; DB 4; Length 433;
Best Local Similarity 86.2%; Pred. No. 2.9e-160;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

```

```

QY 206 NDVARGIVKADVAQSSSYGLGOGQIVAVADTGIDTGRNDSMHEAFRGKITLALYALGRIN 265
DB 1 NDVARGIVKADVAQNNNGYLGOGQIVAVADTGIDTGRNDSMHEAFRGKITLALYALGRIN 60
QY 266 NNDRTGHTHTVAGSYLGGXTNKGMARQANLYFQSIMSXGGLGGLPSNLDTLESQAXS 325
DB 61 NASDPHGHTVAGSYLGN-ALNKGMARQANLYFQSIMSXGGLGGLPSNLDTLESQAXS 119
QY 326 AGARHTNSGMAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNGGTTISAPGTAKNAI 385
DB 120 AGARHTNSGMAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNGGTTISAPGTAKNAI 179
QY 386 TVGATEENLRFSGSYADININHAQFSSRGPTKDGRIKPDVMAFGTITLSARSLADSSR 445
DB 180 TVGATEENLRFSGSIADININHAQFSSRGATRDGRIRKPDVTAAGTITLSARSLADSSR 239
QY 446 NNNHSHKAYMGSTNATITVAGNVAQLREHFKYKNGITPKPSLLKAALILAGADGLGY 505
DB 240 NNNHSHKAYMGSTNATITVAGNVAQLREHFKYKNGITPKPSLLKAALILAGADGLGY 299
QY 506 PANGOGMCHVTLDKSLNVAVYNESSXLTSTOKATYFTATAGPKLISLWSDAPASTTA 565
DB 300 PANGOGMCHVTLDKSLNVAVYNESSXLTSTOKATYFTATAGPKLISLWSDAPASTTA 359
QY 566 SYTLVNDLDTLTAPNGTYVGNDEKXXPKXMMDCGRNNVENFVFNXPOSGTYTIEVOAYN 625
DB 360 SYTLVNDLDTLTAPNGTYVGNDEKXXPKXMMDCGRNNVENFVFNXPOSGTYTIEVOAYN 419
QY 626 VPVGPOXFSLAIVN 639
DB 420 VPVGPOXFSLAIVH 433

```

```

RESULT 10
US-08-894-818B-1
; Sequence 1, Application US/08894818B
; Patent No. 6261822
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masanori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,818B
; FILING DATE: 20-May-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03253
; FILING DATE: 07-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 323285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TAKAKURA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-818B-1

```

```

Query Match          14.6% Score 451.5; DB 4; Length 659;
Best Local Similarity 25.9%; Pred. No. 5.6e-31;
Matches 157; Conservative 92; Mismatches 207; Indels 151; Gaps 24;

```

```

QY 69 LKGLKKLETPYRANKKLIHQFNGLILEETK--QLEXTGAKILDYTFDIYATVIEED 126
DB 43 LTRPGFKKQYRMMNQEVPVTVIMFGSYGDRDRVYKVLRLMGAQY----KYSKI----- 92
QY 127 VXSXXXXIHHVESVEEYLYLXXYIDPOLFTKGASXLVKAAXALDPQKXKEVQLRGIEIXIA 186
DB 93 -----IYAVAK-----IKARDLLIAGMIDTGYRG-NTRVSGIKFIQE 130
QY 187 XXXSNDVXYITAKPEKXVNDVARGI--VKADVAQSSYGLYSGOGQIVAVADTGIDTGRND 244
DB 131 -----DYKVOVDATSVSQCADPTVMNSLGYDSGVVVAIVTDGIDAN--- 173
QY 245 SSMHEAFRGKITALY-ALGRTNNANDPNGHGHVAGSYLGNCKTNK---GMARQANLYPQ 300
DB 174 ---HPDLKGRVIGWYDAVNGRSTPYDDGHHGVHAGIVAGTSVNSQYIGVAPGAKLVG 230
QY 301 SIM--DSXGGLGLPSNLDTLESQAXSAGARI-----HTNSGMAVNGAYTTD 346
DB 231 KYLGADSGSVSTIITAGVWVYQNKKYGIRVINLSLSSQSSDGDLSQAVNNAMWDA- 289
QY 347 SRNVDDYVRKNDMTILFAAGNEXPNGGTTISADGTAKNAITVGATENLRPFSGSYADNINH 406
DB 290 -----GIVVCVAGNSGPNYTVGPSAASKVITVGA-----VDSNDN 327
QY 407 VAOFSRGPTKGRIRKPDVMAFGTITLSARSLADSSRPMANHDSKYAYMGSTNATITV 466
DB 328 IASFSRGPTAQRKPEVAVAPGVDTIAPRA--GTSKGTPIINDYTTASGTSMATPIV 384
QY 467 AGNVADLRHFPVKNRGITP---KPSLLKAA-LIAGAADGLGVPNGOGMCHVTLDKSL- 521
DB 385 SGVGLALID--AHPSTWTDKTKTALITADIAPREINDIAY-----GAGRVNYYKALK 436
QY 522 --NVAYNDSXLTSTOKATYFTATAGPKLISLWSDAPASTASVTLVNDLDTLVTA 579
DB 437 YDDYAKLFTGSAADKGSATHTFEDVSGATFVATLVWD-----TGSSDIDLXLVD 486
QY 580 PNG-----TYVGNDEKXXPKXMMDCGRNNVENFVFNXPOSGTYTIEVOAYNVPVPOX 632
DB 487 PNGNEVDVSYTAYYG-----FEKVGYNPAGTWTWKVVSYK---GAAN 527

```

OY 633 FSLAIYN 639
: : :
DB 528 YQVDVVS 534

RESULT 11
US-09-445-472-12
; Sequence 12, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Kunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/09/445,472
; CURRENT FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-09-445-472-12

Query Match 14.6%; Score 451.5; DB 4; Length 659;
Best Local Similarity 25.9%; Pred. No. 5,6e-31;
Matches 157; Conservative 92; Mismatches 207; Indels 151; Gaps 24;
OY 69 LKRLKLEETVANKKLIHQFNGPILEETK--QXLEXTGAKILDIYIPDIATVYEED 126
DB 43 LTPLEKRVKVRMMNDEVTVMFESYCDRRAVAVLRMLGAOV-----KISTKI----- 92
OY 127 VXSXXXIIEHVESVEPYLPXYXIDPOLFTKGASXLVAKALDTKQNKKEVQLGIEIXIAQ 186
DB 93 -----IPAVAK-----IKARDLLLAGIMIDGYRG-NRVSIGKIFIDE 130
OY 187 XXXSNDVXYITAKPEYKVNDAVARGI--VKADVAOSSYGLYGOGQIVAAVADTGLDTRND 244
DB 131 -----DYKVVDDATSVSIOGADTWNNSIGYDGSVVVAIVDGDIDAN--- 173
OY 245 SSMHEARCKITALY-ALGRTNNANDTNGHGTIVAGSVLGNKXTNK---GMAPOANLVFQ 300
DB 174 ---HPDLKGVIGWYDAVNGRSTPYDDQGHVAGIYAGTGSVNSQYIGVAPGAKLVGV 230
OY 301 SIM--DSXGGLGGLPNSLQTLFSAKXAGART-----HTNSGCAVNGAYTTD 346
DB 231 KVLGADGSGSVSTIIAGVDMVYQNKDKYGIIVNLSSGSSOSSDGIDSLQAVNNAMWA- 289
OY 347 SRNVDDVVRKNDMTLIFAGNEXPNNGTISAPGTAKNAITVGATENRPSFGSYADNINH 406
DB 290 -----GLVYCAAGNSGPTTYVGSPLAASKYITVGA-----VDSNDN 327
OY 407 VAOFSSNGPRTKRIKEDYMAFGTYILSARBSLAPDSSFNAMHDSKATVWGTSMATPIV 466
DB 328 IASSSNGPFAADRKLKPEVAVAGVDIAPRAS--GTSKGTPIVDYTTASGTSMATPIV 384
OY 467 AGNVAQLREHFVKNRGITP--KPSLAKA-LIAGAADXLGLYPNNGQNGRVTLDKSL- 521
DB 385 SGVCAILLQ---AHPSWTDDKVKYKTALIEFADIVAPKEIADIAV---GAGRNVYKAK 436
OY 522 ---NVAVNESSLSQSKATYXFTATAGKPLKISLVSDAPASTTASVTLVNDLVLITA 579
DB 437 YDDYAKLLETFGSAVDKGSATHTFEDVSGATFYATLTYMD-----TGSSPIDLYLTD 486
OY 580 PNG-----TXVYVNDFFXXPYXXNMDCGRNNVENVFINXPOSGTYTIEVOAYVWPGPOX 632
DB 487 PNGNEVDVSYTAYYG-----FEKVGYYNPTAGTWTYKVVSYK---GAAN 527

OY 633 FSLAIYN 639
: : :
DB 528 YQVDVVS 534

RESULT 12
US-08-894-818B-35
; Sequence 35, Application US/08894818B
; Patent No. 6261822
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masanori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Kunoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,818B
; FILING DATE: 20-MAY-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03253
; FILING DATE: 07-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 323285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TAKAKURA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 654 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-818B-35

Query Match 13.2%; Score 408; DB 4; Length 654;
Best Local Similarity 25.9%; Pred. No. 3,3e-27;
Matches 162; Conservative 80; Mismatches 210; Indels 174; Gaps 26;
OY 62 VESENVKLRKGLKXKLETVANKKLIH--IXQFNGPILEETKQXLEXTGAKILDIYIPVAV 119
DB 35 VEKNVGLTLPGLFRKIKLOKLPNBEISTVTFPENHREKEIAVRVLELGAHV-----R 87
OY 120 IVEYEGDVXSXXXIIEHVESVEPYLPXYXIDPOLFTKGASXLVAKALDTKQNKKEVQLR 179
DB 88 V-----YHILIPAI-----AADLKVNDLVLISGLYGRKAKLS 118
OY 180 GLEXIAQXXXSNDVXYITAKPEYKVNDAVARGIVKADVAOSSYGLYGOGQIVAAVADTGLD 239
DB 119 GVRFTQEDYK-----VTVAELEGDESAAQVATVYMNIGYD--GSGITIGITIDTGID 170

QY	240	TGRDSSWHEEFKRTITALVALGCTNNAN-----DTNGHGTIVAGSVLENGXTN-----K	289
Db	171	-----ASHPOLQKRV-----ICWYD-VYNGRSRPIYDDHGHTHVASITAACTGAASNGKYK	219
QY	290	GMAPOANLVFOSIN--DSXGGLGGLPSNLQTLFESQAXSAGARIHTNSWGA-----	337
Db	220	GMAFGARKLAGIKYVLGADGSGSISTIKIGVEWADNKKDYGKIKYINLISGSSQSDGDAL	279
QY	338	-AANGATTDTSRRVDDVYKRNMDTILFPAAGNEKPNCGCTISAPOTAKNATTVGATELNR	395
Db	280	SOAVNAAMDA-----GLVVVVAAGNSGPNKTYTTGSPAAASKVTVGA-----	321
QY	396	SFGSYADININHVAOFSRSKCPKDKIRIKPDYMAPETXLARSLSLAPDSEFWAMHDSKAY	455
Db	322	-----VQKYDVTSEFSSGPTADGRLPPEYVAPENMTIARAS---GTSNGQPIINDYYTA	373
QY	456	MGGTSMATPIYAGNVAOLREHFVYNNRBITER--PSLLKAALLIAGA-----ADGXGLGY	505
Db	374	APGTSMAAPPHVAGTIAALLLO-----AHSWTPDCKVKTLLETADIVKDEIDAD--IAY	424
QY	506	PNGNOGMCGRVTLDRKSLUNAVYNESXLSSTSQKA-----TAXXTATAGKPLKISLWSDAP	560
Db	425	-----GACRVAAKYAIAIN--IDNVAKLVFTGYVANKSGQTHQFYISGASFTYATILYMDAN	477
QY	561	ASTTASVTLVNDLIDLVTTPANG-----TXYVGNDEXXPEXXXNMGDNNVENVFLINXPO	613
Db	478	-----SDLLLYLVDPRNGNOVDYSYTAAYG-----FEKVGYYNPT	511
QY	614	SGTYTIEQAAVNVPEQOXPSLAIYVN	639
Db	512	DGTYTIRIKVAYS---GSANVOVDVVS	534

RESULT 13
 US-09-445-472-16
 Sequence 16, Application US/09445472
 Patent No. 6358726
 GENERAL INFORMATION:
 APPLICANT: TAKAKURA, Hikaru
 APPLICANT: MORISHITA, MIO
 APPLICANT: SHIMOTO, Tomoko
 APPLICANT: ASADA, Kiyozo
 APPLICANT: KAO, Ikunoshin
 TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
 FILE REFERENCE: TAKAKURA-6
 CURRENT APPLICATION NUMBER: US/09/445,472
 CURRENT FILING DATE: 1999-12-06
 PRIOR APPLICATION NUMBER: 151969/1997
 PRIOR FILING DATE: 1997-06-10
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 16
 LENGTH: 654
 TYPE: PRT
 ORGANISM: Pyrococcus furiosus
 US-09-445-472-16

```

Query Match      13.2%; Score 408; DB 4; Length 654;
Best Local Similarity 25.9%; Pred. No. 3.3e-27;
Matches 162; Conservative 80; Mismatches 210; Indels 174; Gaps 26;

QY 62 VESENVKLLXKGLKKLETPVANNKLH--IXQNGPILBETKQXLEXTGAKRIDYIPDAYV 119
      || |||:::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 35 VEKNYGLTLPGLFKRIOQLNPNEISITVIFENRHEKEIARVLELMGAKV-----RY 87

QY 120 IVEYEGDVAXSXXXIEHVESVEPPILPYXIIDPQLFKGASXLYKXAKALDFOKXNKEVQIR 179
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 88 V-----YHIIPAI---AADELVRBDLLVIGSLTGKAKLS 118

QY 180 GLEIXIAOXXXSNDVYITRAKPEYKVMNDVARSIVKADVAQSSYGLGCGQIVAAVDTGID 239
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 GVRFLQGEYK-----VIVSAELGELDSDSAOVAKATVWNIGYD--GSGITTLGIIIDTGID 170

```

QY	240	TGRNDSSWHEAFRCKITALTALVLTGFTNNAN-----DTNGHGTTHVAGSVLGNKXTN-----K	289
Db	171	-----ASHPDLQKV-----IGWVDVYNGSRSTPYDDHGHGTHVASIAAGTGAASNGKY	219
QY	290	GMAPQANLVFOSIM--DSXGGLGGLPSNLQTLFQSAXSAGARIHTNSWGA-----	357
Db	220	GMAPKARLAGIKVLGADSGSISITLIGKEVAVYNNKKRGYKIKVINTLSISSQSDGTDAL	279
QY	338	-----AVNGATTTSSRNVDDYVRKNDMTILFPAAGNEKPNCGTISAPOTANATTVGATENLR	395
Db	280	SOAVNAAMDA-----GLVVVAVVAAAGNSCPNRYTTIGSPAAASKVITVGA-----	321
QY	336	SFGSYADNININVAOFSSRGPKPDRIKIPDWAPGTXLISARSLAPDSSFWAMHDSXYAY	455
Db	332	-----YDKYDVTITSSSGPFIADRLKPEVYAPENMTIARAS---GTSNGQPIINDYYTA	373
QY	456	MCGTSMATPIYAGVAVQULREHFVYNNRGITRK--PSLIKAALILAGA-----ADXGLGY	505
Db	374	APGTSMAHPHVAAGIAAALLQ-----AHPSPWPDKKVKTALFETADIVRKDEIAD--IAY	424
QY	506	PNGNGMGWRVLLDLSINAVYNESSXLSTSOKA-----TXAFTATACKPLKISLWSDAP	560
Db	425	-----GACRVANAYRAIN--YDNVAKLVFTGYVAKKSGQTHQFVYISGASFTATLWMDNAN	477
QY	561	ASTPASVTLVNDLDELVTIAPNG-----TXYVGNDEXXEPXXXXNMWDGRNNVENVFNIXPO	613
Db	478	-----SDLDLXYLDPNGNVDDYSYTAAYG-----FEKVGYYNPT	511
QY	614	SGTTYIEVOQAVNVPQXESLAIYN	639
Db	512	DGTWTKIVSYSS--GSANVQVDVDS	534

RESULT 14
 US-08-894-818B-5, Application US/08894818B
 ; Sequence 5, Application US/08894818B
 ; Patent No. 6261822
 ; GENERAL INFORMATION:
 ; APPLICANT: TAKAKURA, Hikaru
 ; APPLICANT: MORISHITA, Mio
 ; APPLICANT: YAMAMOTO, Katsuniko
 ; APPLICANT: MITTA, Masaoori
 ; APPLICANT: ASADA, Kiyozo
 ; APPLICANT: TSUNASAWA, Susumu
 ; APPLICANT: KATO, Ikunoshin
 ; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Browdy and Neimark
 ; STREET: 419 Seventh Street N.W., Ste. 300
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: United States of America
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/894,818B
 ; FILING DATE: 20-MAY-1998
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/JP96/03253
 ; FILING DATE: 07-NOV-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 33285/1995
 ; FILING DATE: 12-DEC-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Browdy, Roger L.
 ; REGISTRATION NUMBER: 25,618


```
REFERENCE/DOCKET NUMBER: TAKAKURA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-894-818B-5
```

```
Query Match      12.8%: Score 398; DB 4: Length 659;
Best Local Similarity 24.4%: Pred. No. 2.4e-26;
Matches 151; Conservative 94; Mismatches 215; Indels 158; Gaps 24;
```

```
OY 62 VESNNVAKLKKLLETPANNKLIH--IXQFNGPILEETKQXLEXTGAKIIDIYIPDAY 119
DB 35 VEKNYGLLPGLERKIQKLNPEEISTYIVFENHREKEIAVRLELMGAKV-----RY 87
OY 120 IVEYEGDVSKXXXIEHVESVEPYLPXYXIDPOLFTKGSXIVKAXALDPTKQNKVEQLR 179
DB 88 V-----YHIIPAI--ADLKRDLVLSGLTGKAKLS 118
OY 180 GIEIXAQXXXNDVYITAKPEYKMDVARGIVKADVAQSSYGLYGQOIYAADTGID 239
DB 119 GVRFIQEDYK-----VTVSALEGLDESAAQVATYVWMLGYD--GSGITIGIDTGD 170
OY 240 TGRNDSMHEAFRGKITLALYALGRTNNAN-----DTNGHGFHVAGSVLNGXTN---K 289
DB 171 -----ASHPDLOGKV-----IGWDFVNGRSYPYDDHGHGHVSAIAGTGAASNGKRYK 219
OY 290 GMAPQANLVFQSIM--DSXGGLGGLPSNLQTLFSQAXSAGARI-----HTNSW 335
DB 220 GMAPGALAGIKVLAGDSGSIITIKGEVAVDNKDKGIKVINLSGSSOSSDGTSL 279
OY 336 GAAVNGAYTTDSRNNDDVYRKNDMTILEFAGENEXPNGTISAPGTAKNATYVATENLRP 395
DB 280 SQAVNNAMDA-----GLVVCVAAGNSGPNYTVVGSFAASKRYTVGA----- 321
OY 396 SFGSYADININHAQFSSRGPTKGRIRPDVMAPOGTXILSARSLADSSFMANHDSKYAY 455
DB 322 -----VDSNDNIASFSSRGPTADGRLKREYVAPGVLDIAPRAS--GTSMTGPIINDYTK 373
OY 456 MGTSMATPIVAGNVAQLREHFVKNRGITP--KPSILKAA--LIAGADXGLGYPNGNOG 511
DB 374 ASGTSMATPIVSGVGLALIQ--APSWTPDKVYKTALIEADIVAPKEIADIAY-----G 425
OY 512 WGRVTTLDKSL--NVAAYNESXSLSTSOKATYXFTATAGKPLKISLWMSDAPASTASYT 568
DB 426 AGRVNVYKAIKYDDYAKLFTGVSADKGSATHFEDVSGATFVATLYMD-----T 475
OY 569 LVNDLDELVTAPNG-----TXVYGNDFKXPRXXNMMDGNNEVFINXPOSGTYTEV 621
DB 476 GSSDIDLXLDPNGNEVDYSTAYTG-----FEKVGYYNPATGATWTYKV 519
OY 622 QAYNVPGQXFSLAIVN 639
DB 520 VSTK--GAANTQVDVVS 534
```

```
RESULT 15
US-09-445-472-1
; Sequence 1, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshi
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
```

```
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/09/445,472
CURRENT FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 412
TYPE: PRT
ORGANISM: Pyrococcus furiosus
US-09-445-472-1
```

```
Query Match      12.6%: Score 391; DB 4: Length 412;
Best Local Similarity 28.9%: Pred. No. 5.2e-26;
Matches 134; Conservative 56; Mismatches 145; Indels 128; Gaps 20;
```

```
OY 223 GLYGQOIYAAVADGDLTGRNDSMHEAFRGKITLALYALGRTNNAN-----DTNGHGH 276
DB 22 GYDGSGITIGIDTGDID-----ASHPDLOGKV-----IGWDFVNGRSYPYDDHGHGH 70
OY 277 VAGSVLNGXTN--KGMAPQANLVFQSIM--DSXGGLGGLPSNLQTLFSQAXSAGARI 330
DB 71 VASIAAGTGAASNGKRYKGMAPGAKLAGIKVLAGDSGSIITIKGEVAVDNKDKYGIIV 130
OY 331 HTNSMGA-----AVNGAYTTDSRNDVYRKNDMTILEFAGENEXPNGTISAP 378
DB 131 INLSGSSOSSDGTDLALSQAVNMAA-----GLVVCVAAGNSGPNKTYTIGSP 178
OY 379 GTAKNATYVATENLRPSFGSYADININHAQFSSRGPTKGRIRPDVMAPOGTXILSARSS 438
DB 179 AASKRYTVGA-----VDKVDVITSFSSRGPTADGRLKREYVAPGVLDIAPRAS 227
OY 439 LAPDSFMANHDSKAIYAGTSMATPIVAGNVAQLREHFVKNRGITPK--PILKALIA 496
DB 228 --GTSMGOPINDYTTAAPGTSMATPIVAGIALLLQ-----AHPSWTPDKVKTALIE 277
OY 497 GA-----ADXGLGYPNGNGRVTLDKSLNVAAYNESXSLSTSOKA-----TYXFT 543
DB 278 TADIVKPEIAD--IAY-----GAGRVAAYKAIN--YDVTAKLVFTGIVANKGSOTHOEV 328
OY 544 ATAGKPLKISLWMSDAPASTASYTLVNDLDELVTAPNG-----TXVYGNDFKXPRXX 596
DB 329 ISGASFVATILYMDNAN-----SDLILYLDLPNGNOVDYSTAYYG----- 369
OY 597 NMDGRNNVENVFINXPOSGTYTEVQAYNVPGQXFSLAIVN 639
DB 370 -----FEKVGYYNPATGATWTIYVVSYS--GSANYQVDVVS 402
```

```
Search completed: April 1, 2003, 12:42:09
Job time : 17.9891 secs
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: April 1, 2003, 12:41:04 ; Search time 23.9812 Seconds
(without alignments)
1629.021 Million cell updates/sec

Title: US-09-920-954-1
Percent score: 3101
Sequence: 1 MKKKVFLSVLSAAAILSTV.....EVOAYNPVGFQXETSLAINV 639

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA.*
1: /cgn2_6/ptodata/1/pubppa/US08_NEM_PUB.pep.*
2: /cgn2_6/ptodata/1/pubppa/PCIT_NEM_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_NEM_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_NEM_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppa/PCITUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_NEM_PUB.pep.*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppa/US10_NEM_PUB.pep.*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubppa/US60_NEM_PUB.pep.*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2155	69.5	434	9	US-09-985-689A-1
2	2155	69.5	434	9	US-09-985-689A-2
3	2082	67.1	434	9	US-09-985-689A-6
4	2060.5	66.4	433	9	US-09-985-689A-7
5	1952.5	63.0	433	9	US-09-985-689A-5
6	1948.5	62.8	433	9	US-09-985-689A-3
7	1941.5	62.6	433	9	US-09-985-689A-4
8	451.5	14.6	659	12	US-10-090-624-12
9	408	13.2	654	12	US-10-090-624-16
10	391	12.6	522	12	US-10-090-624-1
11	391	12.6	522	12	US-10-090-624-11
12	307.5	9.9	418	10	US-09-966-921A-2
13	289.5	9.7	1398	12	US-10-090-624-6
14	280.5	9.0	595	9	US-09-927-827-59
15	271.5	8.8	580	9	US-09-927-827-55
16	270.5	8.7	379	9	US-09-813-408-6
17	270.5	8.7	397	9	US-09-779-334A-5
18	250.5	8.2	382	12	US-10-090-624-31
19	250.5	8.1	380	9	US-09-813-408-16

20	250.5	8.1	380	9	US-09-824-893A-261	Sequence 261, App
21	250	8.1	377	9	US-09-813-408-1	Sequence 1, Appl
22	250	8.1	379	9	US-09-813-408-13	Sequence 13, Appl
23	250	8.1	380	9	US-09-813-408-17	Sequence 19, Appl
24	250	8.1	382	9	US-09-813-408-10	Sequence 7, Appl
25	249.5	8.0	379	9	US-09-813-408-10	Sequence 10, Appl
26	246	7.9	271	9	US-09-813-408-2	Sequence 11, Appl
27	244.5	7.9	379	9	US-09-813-408-11	Sequence 16, Appl
28	244.5	7.9	381	10	US-09-920-118-16	Sequence 14, Appl
29	242.5	7.8	379	10	US-09-920-118-14	Sequence 1, Appl
30	239	7.7	269	9	US-09-975-139-1	Sequence 8, Appl
31	239	7.7	269	9	US-09-976-414-8	Sequence 16, Appl
32	239	7.7	269	10	US-09-837-235-16	Sequence 6, Appl
33	239	7.7	269	10	US-09-060-8548-6	Sequence 1, Appl
34	239	7.7	269	12	US-10-075-907-1	Sequence 1, Appl
35	239	7.7	269	12	US-10-075-895-1	Sequence 4, Appl
36	235	7.6	269	9	US-09-813-408-4	Sequence 14, Appl
37	235	7.6	378	9	US-09-813-408-14	Sequence 3, Appl
38	234.5	7.6	1150	10	US-09-870-122-3	Sequence 23, Appl
39	234.5	7.6	1181	10	US-09-870-122-23	Sequence 15, Appl
40	234	7.5	310	9	US-09-813-408-15	Sequence 5, Appl
41	232	7.5	269	9	US-09-813-408-5	Sequence 21, Appl
42	231.5	7.5	1167	10	US-09-870-122-2	Sequence 9, Appl
43	230.5	7.4	279	9	US-09-813-408-21	Sequence 7, Appl
44	229.5	7.4	372	9	US-09-813-408-9	
45	229	7.4	269	9	US-09-779-334A-7	

ALIGNMENTS

RESULT 1
US-09-985-689A-1
; Sequence 1, Application US/09985689A
; Publication NO. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUOSHI
; APPLICANT: SAKETI, KATSUHIKA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985.689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR FILING DATE: 2000-11-22
; PRIOR FILING DATE: 2000-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-1
Query Match 69.5% Score 2155 DB 9 Length 434
Best local Similarity 96.3% Pred No 3 9e-175
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
OY 206 NDVAKRIVADVAQSSYGLKQGGIVAVADTGLDGRNNSMHEAFKRIKITALVALGRTN 255
DB 1 NDVAKRIVADVAQSSYGLKQGGIVAVADTGLDGRNNSMHEAFKRIKITALVALGRTN 60
OY 266 NADNTGSGHTVAGSVLNGKTYNKGMAPQANLVFQSIMDSXGGLGSLPNSLQTLFQSAXS 335
DB 61 NADNTGSGHTVAGSVLNGKTYNKGMAPQANLVFQSIMDSXGGLGSLPNSLQTLFQSAXS 120
OY 326 AGARITNSMGAANGAVTTDSRVVDYVRKNDMTILFAAGNEXPNGCTISAPCTAKNAI 385

```
|||||
Db 121 ACARHTNSWGAAYVAGTDDSRNVDYRKNDMTLFAAGNEGPGGTISAPGAKNAI 180
Qy 386 TVGATENLRPSFGSYADNINNHVAFSSRGPTRDKRIKPDVMAPGTYILSARSLADDSF 445
Db 181 TVGATENLRPSFGSYADNINNHVAFSSRGPTRDKRIKPDVMAPGTYILSARSLADDSF 240
Qy 446 WANHDSKYAYMGTSMAPIVAGNVAQLREHFVKNGITPKPSLLKALILAGAADXGLGY 505
Db 241 WANHDSKYAYMGTSMAPIVAGNVAQLREHFVKNGITPKPSLLKALILAGAADXGLGY 300
Qy 506 PNGNOGWRVTLDKSLNVAAYVNESSSLSTQKATYFTATACKPLKISLWSDAPASTTA 565
Db 301 PNGNOGWRVTLDKSLNVAAYVNESSSLSTQKATYFTATACKPLKISLWSDAPASTTA 360
Qy 566 SVTLVNDLDTLTAPNGITVYVGNDFXPXXMMWDCRNNEVFINXPOSGTITIEVOAYN 625
Db 361 SVTLVNDLDTLTAPNGITVYVGNDFXPXXMMWDCRNNEVFINXPOSGTITIEVOAYN 420
Qy 626 VPGPQXFSLATVN 639
Db 421 VPGPQXFSLATVN 434
```

```
RESULT 2
: Sequence 2, Application US/09985689A
: Publication No. US20030022351A1
: GENERAL INFORMATION:
: APPLICANT: HATADA, YUJI
: APPLICANT: OGAWA, AKINORI
: APPLICANT: KAGEYAMA, YASUSHI
: APPLICANT: SATO, TSUYOSHI
: APPLICANT: ARAKI, HIROYUKI
: APPLICANT: SUMITOMO, NOBUYUKI
: APPLICANT: OKUDA, MITSUYOSHI
: APPLICANT: SAEKI, KATSUHIISA
: TITLE OF INVENTION: Alkaline proteases
: FILE REFERENCE: 215483050
: CURRENT APPLICATION NUMBER: US/09/985,689A
: PRIOR FILING DATE: 2002-07-01
: PRIOR APPLICATION NUMBER: JP P2000-355166
: PRIOR FILING DATE: 2000-11-22
: PRIOR APPLICATION NUMBER: JP P2001-114048
: PRIOR FILING DATE: 2001-04-12
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 2
: LENGTH: 434
: TYPE: PRT
: ORGANISM: Bacillus sp.
US-09-985-689A-2
```

Query Match 69.5%; Score 2155; DB 9; Length 434;

Best Local Similarity 96.3%; Pred. No. 3.9e-175; Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy 206 NDVARGIYKADYAOSSYGLYGOGIYVAVDTGLDGRNDSMHEAFRGKITLALYLGRTN 265
Db 1 NDVARGIYKADYAOSSYGLYGOGIYVAVDTGLDGRNDSMHEAFRGKITLALYLGRTN 60
Qy 266 NANDNGHGTTHVAGSVLNGXTNKGMAPOANLVFOSIMDSXGGLGGLPSNLQTLFSQAYS 325
Db 61 NANDNGHGTTHVAGSVLNGXTNKGMAPOANLVFOSIMDSXGGLGGLPSNLQTLFSQAYS 120
Qy 326 AGARHTNSWGAAYVAGTDDSRNVDYRKNDMTLFAAGNEGPGGTISAPGAKNAI 385
Db 121 AGARHTNSWGAAYVAGTDDSRNVDYRKNDMTLFAAGNEGPGGTISAPGAKNAI 180
Qy 386 TVGATENLRPSFGSYADNINNHVAFSSRGPTRDKRIKPDVMAPGTYILSARSLADDSF 445
Db 181 TVGATENLRPSFGSYADNINNHVAFSSRGPTRDKRIKPDVMAPGTYILSARSLADDSF 240
```

```
Qy 446 WANHDSKYAYMGTSMAPIVAGNVAQLREHFVKNGITPKPSLLKALILAGAADXGLGY 505
Db 241 WANHDSKYAYMGTSMAPIVAGNVAQLREHFVKNGITPKPSLLKALILAGAADXGLGY 300
Qy 506 PNGNOGWRVTLDKSLNVAAYVNESSSLSTQKATYFTATACKPLKISLWSDAPASTTA 565
Db 301 PNGNOGWRVTLDKSLNVAAYVNESSSLSTQKATYFTATACKPLKISLWSDAPASTTA 360
Qy 566 SVTLVNDLDTLTAPNGITVYVGNDFXPXXMMWDCRNNEVFINXPOSGTITIEVOAYN 625
Db 361 SVTLVNDLDTLTAPNGITVYVGNDFXPXXMMWDCRNNEVFINXPOSGTITIEVOAYN 420
Qy 626 VPGPQXFSLATVN 639
Db 421 VPGPQXFSLATVN 434
```

```
RESULT 3
: Sequence 6, Application US/09985689A
: Publication No. US20030022351A1
: GENERAL INFORMATION:
: APPLICANT: HATADA, YUJI
: APPLICANT: OGAWA, AKINORI
: APPLICANT: KAGEYAMA, YASUSHI
: APPLICANT: SATO, TSUYOSHI
: APPLICANT: ARAKI, HIROYUKI
: APPLICANT: SUMITOMO, NOBUYUKI
: APPLICANT: OKUDA, MITSUYOSHI
: APPLICANT: SAEKI, KATSUHIISA
: TITLE OF INVENTION: Alkaline proteases
: FILE REFERENCE: 215483050
: CURRENT APPLICATION NUMBER: US/09/985,689A
: PRIOR FILING DATE: 2002-07-01
: PRIOR APPLICATION NUMBER: JP P2000-355166
: PRIOR FILING DATE: 2000-11-22
: PRIOR APPLICATION NUMBER: JP P2001-114048
: PRIOR FILING DATE: 2001-04-12
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 6
: LENGTH: 434
: TYPE: PRT
: ORGANISM: Bacillus sp.
US-09-985-689A-6
```

Query Match 67.1%; Score 2082; DB 9; Length 434;

Best Local Similarity 91.5%; Pred. No. 6.2e-169; Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

```
Qy 206 NDVARGIYKADYAOSSYGLYGOGIYVAVDTGLDGRNDSMHEAFRGKITLALYLGRTN 265
Db 1 NDVARGIYKADYAOSSYGLYGOGIYVAVDTGLDGRNDSMHEAFRGKITLALYLGRTN 60
Qy 266 NANDNGHGTTHVAGSVLNGXTNKGMAPOANLVFOSIMDSXGGLGGLPSNLQTLFSQAYS 325
Db 61 NANDNGHGTTHVAGSVLNGXTNKGMAPOANLVFOSIMDSXGGLGGLPSNLQTLFSQAYS 120
Qy 326 AGARHTNSWGAAYVAGTDDSRNVDYRKNDMTLFAAGNEGPGGTISAPGAKNAI 385
Db 121 AGARHTNSWGAAYVAGTDDSRNVDYRKNDMTLFAAGNEGPGGTISAPGAKNAI 180
Qy 386 TVGATENLRPSFGSYADNINNHVAFSSRGPTRDKRIKPDVMAPGTYILSARSLADDSF 445
Db 181 TVGATENLRPSFGSYADNINNHVAFSSRGPTRDKRIKPDVMAPGTYILSARSLADDSF 240
Qy 446 WANHDSKYAYMGTSMAPIVAGNVAQLREHFVKNGITPKPSLLKALILAGAADXGLGY 505
Db 241 WANHDSKYAYMGTSMAPIVAGNVAQLREHFVKNGITPKPSLLKALILAGAADXGLGY 300
Qy 506 PNGNOGWRVTLDKSLNVAAYVNESSSLSTQKATYFTATACKPLKISLWSDAPASTTA 565
Db 301 PNGNOGWRVTLDKSLNVAAYVNESSSLSTQKATYFTATACKPLKISLWSDAPASTTA 360
```

```

Qy 566 SVTLVNDLDTLTAPNGTXYVGNDFXXPXXXNMDGRNNVENVFIMXPOSGTYTTEVOAYN 625
Db 361 SVTLVNDLDTLTAPNGTXYVGNDFXAPYDNNMDGRNNVENVFIMXPOSGTYTTEVOAYN 420
Qy 626 VPGVPOXFSLAIVN 639
Db 421 VPGVPOXFSLAIVN 434

RESULT 4
US-09-985-689A-7
; Sequence 7, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAMA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483050
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 433
; TYPE: prt
; ORGANISM: Bacillus sp.
US-09-985-689A-7

Query Match 66.4%; Score 2060.5; DB 9; Length 433;
Best Local Similarity 91.7%; Pred. No. 4,1e-167;
Matches 398; Conservative 12; Mismatches 23; Indels 1; Gaps 1;

Qy 206 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTGRNDSMHEAFRGKITATYALGRTN 265
Db 1 NDVARGIVADVAQNNFGLYGQGIYAVADTGLDTGRNDSMHEAFRGKITATYALGRTN 60
Qy 266 NANDTNGHGTIVAGSVLGNXTNKGMAPQANLVFQSIMDSXGGLGRLPSNLQTLFSGAXS 325
Db 61 NANDPBGHGTIVAGSVLGN-ATNKGMAPQANLVFQSIMDSXGGLGRLPSNLQTLFSGAXS 119
Qy 326 AGARLHTNSMGAAYNGAVTTDSRNVDYRKNDMTILFAAGNEXPNNGTISAPGTAKNAI 385
Db 120 AGARLHTNSMGAAYNGAVTTDSRNVDYRKNDMTILFAAGNEXPNNGTISAPGTAKNAI 179
Qy 386 TYGATENLRPSFGSYADNINHVAFQSSRGPTKGRKPDVMAFGTYILSARSSLAPDSSF 445
Db 180 TYGATENLRPSFGSYADNINHVAFQSSRGPTKGRKPDVMAFGTYILSARSSLAPDSSF 239
Qy 446 WANHDSKYAYMGTSMATPIVAGNVAOQLREHFVKNRGITPKPSLLKALILGADAXGLGY 505
Db 240 WANHDSKYAYMGTSMATPIVAGNVAOQLREHFVKNRGITPKPSLLKALILGADAXGLGY 299
Qy 506 PNGNCGMGRVTLDSKSLNVAAYVNESSXLSTOKATYXFTATAGKPLKISLVWSDAPASTTA 565
Db 300 PNGNCGMGRVTLDSKSLNVAAYVNESSXLSTOKATYXFTATAGKPLKISLVWSDAPASTTA 359
Qy 566 SVTLVNDLDTLTAPNGTXYVGNDFXXPXXXNMDGRNNVENVFIMXPOSGTYTTEVOAYN 625
Db 360 SVTLVNDLDTLTAPNGTXYVGNDFXAPYDNNMDGRNNVENVFIMXPOSGTYTTEVOAYN 419
Qy 626 VPGVPOXFSLAIVN 639
Db 421 VPGVPOXFSLAIVN 434
```

```

Db 420 VPGVPOXFSLAIVH 433

RESULT 5
US-09-985-689A-5
; Sequence 5, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAMA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483050
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 433
; TYPE: prt
; ORGANISM: Bacillus sp.
US-09-985-689A-5

Query Match 63.0%; Score 1952.5; DB 9; Length 433;
Best Local Similarity 86.0%; Pred. No. 6,1e-156;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

Qy 206 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTGRNDSMHEAFRGKITATYALGRTN 265
Db 1 NDVARGIVADVAQNNFGLYGQGIYAVADTGLDTGRNDSMHEAFRGKITATYALGRTN 60
Qy 266 NANDTNGHGTIVAGSVLGNXTNKGMAPQANLVFQSIMDSXGGLGRLPSNLQTLFSGAXS 325
Db 61 NANDPBGHGTIVAGSVLGN-ALNKGMAPQANLVFQSIMDSXGGLGRLPSNLQTLFSGAXS 119
Qy 326 AGARLHTNSMGAAYNGAVTTDSRNVDYRKNDMTILFAAGNEXPNNGTISAPGTAKNAI 385
Db 120 AGARLHTNSMGAAYNGAVTTANSRVDYRKNDMTILFAAGNEXPNNGTISAPGTAKNAI 179
Qy 386 TYGATENLRPSFGSYADNINHVAFQSSRGPTKGRKPDVMAFGTYILSARSSLAPDSSF 445
Db 180 TYGATENLRPSFGSLADPNHIAQFSSRGATRGKRPDVTAPEGFTLSARSSLAPDSSF 239
Qy 446 WANHDSKYAYMGTSMATPIVAGNVAOQLREHFVKNRGITPKPSLLKALILGADAXGLGY 505
Db 240 WANYSKYAYMGTSMATPIVAGNVAOQLREHFVKNRGITPKPSLLKALILGATVDGLGY 299
Qy 506 PNGNCGMGRVTLDSKSLNVAAYVNESSXLSTOKATYXFTATAGKPLKISLVWSDAPASTTA 565
Db 300 PNGNCGMGRVTLDSKSLNVAAYVNEATLATGOKATYSFOAGKPLKISLVWTDAPGSTTA 359
Qy 566 SVTLVNDLDTLTAPNGTXYVGNDFXXPXXXNMDGRNNVENVFIMXPOSGTYTTEVOAYN 625
Db 360 SVTLVNDLDTLTAPNGQKYVGNDFXAPYDNNMDGRNNVENVFIMXPOSGTYTTEVOAYN 419
Qy 626 VPGVPOXFSLAIVN 639
Db 420 VPGVPOXFSLAIVH 433

RESULT 6
US-09-985-689A-3
; Sequence 3, Application US/09985689A
; Publication No. US20030022351A1
```

```
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGWAMA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483050
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-3
```

```
Query Match          62.8%; Score 1948.5; DB 9; Length 433;
Best Local Similarity 86.4%; Pred. No. 1.3e-157;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;
```

```
QY 206 NDVARGIVKADVAOSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 265
D 1 NDVARGIVKADVAONNLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 266 NADPNGHGTIVAGSVLGNKMGAPQANLVFQSIMDSXGGLGGLPSNLTLEFSQAXS 325
D 61 NADPNGHGTIVAGSVLGN -ALNKGMAPQANLVFQSIMDSXGGLGGLPSNLTLEFSQAXN 119
QY 326 AGARLHTSMGAAVNGAYTTDSRVNDVYRKNDMTILFAAGNEXPNCGTISAPGTAKNAI 385
D 120 AGARLHTSMGAPVNGAYTANSRQYDEVYRNNDMTVLFAAGNEGNSGTISAPGTAKNAI 179
QY 386 TVGATENLRPSFGSYADININVAQFSSRGPTKDGRIKPDVAPGTXILSASSLAPDSSF 445
D 180 TVGATENLRPSFGSIADPNHIAQFSSRGATRDGRIKPDVAPGTXILSASSLAPDSSF 239
QY 446 WANDSKYAYMGSTMATPIYAGNVAOLREHFVKNRGTTPEPSLKALALIGAADXGLGY 505
D 240 WANDSKYAYMGSTMATPIYAGNVAOLREHFVKNRGTTPEPSLKALALIGAADXGLGY 299
QY 506 PNGQMGWRTLDSLVNAYVNESSXLSSTQKATYXTATAGKPLKISLVNSDAPASTTA 565
D 300 PNGQMGWRTLDSLVNAYVNEATALTGQKATYSFOTQAGKPLKISLVNTDAGSTTA 359
QY 566 SVTLVNDLDVITAPNGXHYVGNDFXAPXXXNMGRNNEVFNXPQSGTYTTEVQAYN 625
D 360 SVTLVNDLDVITAPNGXHYVGNDFXAPXXXNMGRNNEVFNXPQSGTYTTEVQAYN 419
QY 626 VPVGQFQSLAIYN 639
D 420 VPVGQFQSLAIYH 433
```

```
RESULT 7
US-09-985-689A-4
; Sequence 4, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGWAMA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
```

```
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483050
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-4
```

```
Query Match          62.6%; Score 1941.5; DB 9; Length 433;
Best Local Similarity 86.2%; Pred. No. 5.3e-157;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;
```

```
QY 206 NDVARGIVKADVAOSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 265
D 1 NDVARGIVKADVAONNLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 266 NADPNGHGTIVAGSVLGNKMGAPQANLVFQSIMDSXGGLGGLPSNLTLEFSQAXS 325
D 61 NADPNGHGTIVAGSVLGN -ALNKGMAPQANLVFQSIMDSXGGLGGLPSNLTLEFSQAXN 119
QY 326 AGARLHTSMGAAVNGAYTTDSRVNDVYRKNDMTILFAAGNEXPNCGTISAPGTAKNAI 385
D 120 AGARLHTSMGAPVNGAYTANSRQYDEVYRNNDMTVLFAAGNEGNSGTISAPGTAKNAI 179
QY 386 TVGATENLRPSFGSYADININVAQFSSRGPTKDGRIKPDVAPGTXILSASSLAPDSSF 445
D 180 TVGATENLRPSFGSIADPNHIAQFSSRGATRDGRIKPDVAPGTXILSASSLAPDSSF 239
QY 446 WANDSKYAYMGSTMATPIYAGNVAOLREHFVKNRGTTPEPSLKALALIGAADXGLGY 505
D 240 WANDSKYAYMGSTMATPIYAGNVAOLREHFVKNRGTTPEPSLKALALIGAADXGLGY 299
QY 506 PNGQMGWRTLDSLVNAYVNESSXLSSTQKATYXTATAGKPLKISLVNSDAPASTTA 565
D 300 PNGQMGWRTLDSLVNAYVNEATALTGQKATYSFOTQAGKPLKISLVNTDAGSTTA 359
QY 566 SVTLVNDLDVITAPNGXHYVGNDFXAPXXXNMGRNNEVFNXPQSGTYTTEVQAYN 625
D 360 SVTLVNDLDVITAPNGXHYVGNDFXAPXXXNMGRNNEVFNXPQSGTYTTEVQAYN 419
QY 626 VPVGQFQSLAIYN 639
D 420 VPVGQFQSLAIYH 433
```

```
RESULT 8
US-10-090-624-12
; Sequence 12, Application US/10090624
; Patent No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
```

NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 659
TYPE: PR
ORGANISM: Thermococcus celer
US-10-090-624-12

Query Match 14.6%; Score 451.5; DB 12; Length 659;
Best Local Similarity 25.9%; Pred. No. 3.4e-10;
Matches 157; Conservative 92; Mismatches 207; Indels 151; Gaps 24;

QY 69 LKGLAKKLETVANKKLIHOFNPILEERK--QXLETKGAILDIYIDYAYIEYEGD 126
DB 43 LTPLEFKKQVRMMNOEDVTIMFSSYDGRRAVKVRLMGAVO-----KYSIKI----- 92
QY 127 VXSXXXXIHEVESVEPYLPYXXIDPQLEFTKGSXLYVAKALDTRKXNKEVQIGIEIXAQ 186
DB 93 -----IPAVAVK-----IKARDLLIAGMIDIDYEGE-MTRVSGIKFIQE 130
QY 187 XXXSNDVXYITAKPEYKVNNDVARGI--VKADVAOSSYGLYGOGQIVAVADTGLDTRND 244
DB 131 -----DYKVQVDDATSVQIGADVTWNSLGYDGSVVVAIVDIDGIDAN-- 173
QY 245 SSMHEAFRGKITALY-ALGRTNNANDTNGHGTIVAGSVLNGXTNK---GMAPQANLYEQ 300
DB 174 ---HPDLKGVIGWDAVNGRSTPYDDGCHGTIVAGTGSVNGQYIGVAPGAKIVGV 230
QY 301 SIM--DSXGGLGGLPSNLQTLFSAQXSAGARI-----HTNSGCAVNGCAVTTD 346
DB 231 KVLGAOSSGSVSTIINGVDMVYQNKDYGITIVILSLGSSQSSDGDITSLQAVANNKMDA- 289
QY 347 SRNVDYVRKNDMTILFAAGNEXPNCGTISAPGTAKNAITVGAETENLRPSFGYADNINH 406
DB 290 -----GLVVVAAGNSGPNVTYVGSPPAASKVITVGA-----VDSNDN 327
QY 407 VAQSSSGPTKDGRIKPDVMAFPTXILSARSSLAPDSSFPANHDSKYAYMGTSNATPIV 466
DB 328 IASSSSSGPTPADRLKPEVVAAGVDIIPRAS---GTSMTGPIINDIYATTSAGSMATPIV 384
QY 467 AGNVAQLREHFVKNRGITP--KPSLKA--LIAGADKSLGYPNGNCGMGRVTRDKSL- 521
DB 385 SCVQALTLQ---AHPMTDPDKYKTLIETADIVAPKEIADIAV---GAGRNVYKAK 436
QY 522 --NVAAYNESXSLSTOKATYFTATAGKPLKISLVMSDAPASTFASVTLVNDLDTITA 579
DB 437 YDDYAKLTFGVSADKGSATHTFEDVSGATFVATLLYMD-----TGSSDIDLXYD 486
QY 580 PNG-----TXYVGNDFXPPXXNMWDRNNVENVFINXPOSGTYTIEVOAVNVPRGPX 632
DB 487 PNGNEVDYSTAYYG-----FEKVGYYNPTAGTWTYKVVYSK---GAAN 527
QY 633 FSLAIYN 639
DB 528 XQVDVVS 534

RESULT 9
US-10-090-624-16
Sequence 16, Application US/10090624
Patent No. US20020132335A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOJO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/090,624
PRIOR FILING DATE: 1999-12-06

PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 16
LENGTH: 654
TYPE: PR
ORGANISM: Pyrococcus furiosus
US-10-090-624-16

Query Match 13.2%; Score 408; DB 12; Length 654;
Best Local Similarity 25.9%; Pred. No. 1.7e-26;
Matches 162; Conservative 80; Mismatches 210; Indels 174; Gaps 26;

QY 62 VESENVKLKGLKLETVANKKLIH--IXOFNPILEETKQXLETKGAILDIYIDYAV 119
DB 35 VEKNVGLTLPGLFPRKLOKLPNEISTVIFEFNHEKRIAVRLEMGAV-----RY 87
QY 120 IVEYEGDVXSXXXXIHEVESVEPYLPYXXIDPQLEFTKGSXLYVAKALDTRKXNKEVQLR 179
DB 88 V-----YHILPAI-----AADLKVRLDLYISGLTGKRAKLS 118
QY 180 GLEXIAQXXXSNDVXYITAKPEYKVNNDVARGIVKADVAOSSYGLYGOGQIVAVADTGLD 239
DB 119 GVRFLQEDYK-----VTYSALEGLDESAQVAVTVMNIGVD--GSGITIGIITDIDG 170
QY 240 TGRNDSMHEAFRGKITALYALGRTNNAN-----DTNGHGTIVAGSVLNGXTN-----K 289
DB 171 -----ASHPDLQKV-----IOWDFVNGRSTPYDDHGTHTVAGTGAASNGKK 219
QY 290 GMAPQANLYEQSIM--DSXGGLGGLPSNLQTLFSAQXSAGARIHTNSWGA----- 337
DB 220 GMAPQAKIAGIKVLGADSGSISTITIKGVMAVDNKKDYGITIVILSLGSSQSSDGDIDL 279
QY 338 --AVNGATTTDSRNVDYVRKNDMTILFAAGNEXPNCGTISAPGTAKNAITVGAETENLRP 395
DB 280 SQAVVAAMDA-----GLVVVAAGNSGPNKTYTISPPAASKVITVGA----- 321
QY 336 SFGSYADNINHVAQSSSGPTKDGRIKPDVMAFPTXILSARSSLAPDSSFPANHDSKYAV 455
DB 322 -----VKKYDVTISFSSRGPTADGRKPEVVAAGNMIIAARAS---GTSMGQPIINDIYATTA 373
QY 456 MGTSMATPIYAGNVQDLREHFVKNRGITP--KPSLKAALIAA-----ADXGLCY 505
DB 374 APGTSMATPIYAGIADLLD-----AHPSWTPDKYKTLIETADIVAPKEIADIAV---IAY 424
QY 506 PNGNCGMGRVTRDKSLVAVVNESXSLSTOKA-----TXYFTATAGKPLKISLVMSDAP 560
DB 425 -----GAGRVAATKALIN--YDNVAKLFTGYVANKGSOTHFVLSGASFVATLLYMDNAN 477
QY 561 ASTASVTLVNDLDTYITAPNG-----TXYVGNDFXPPXXNMWDRNNVENVFINXPO 613
DB 478 -----SDDLXYLPNGNOVDYSTAYYG-----FEKVGYYNPT 511
QY 614 SGTITIEYQAVINPVGPQXSLAIYN 639
DB 512 DGTWTKVYSY--GSANYQVDVVS 534

RESULT 10
US-10-090-624-1
Sequence 1, Application US/10090624
Patent No. US20020132335A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOJO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/090,624
PRIOR FILING DATE: 2002-03-06

```

11 RESULT 11
12 US-10-090-624-4
13 ; Sequence 4, Application US/10090624
14 ; Patent No. US2002013235A1
15 ;
16 ; GENERAL INFORMATION:
17 ;
18 ; APPLICANT: TAKAKURA, Hikaru
19 ; APPLICANT: MORISHITA, Mio
20 ; APPLICANT: SHIMOJO, Tomoko
21 ; APPLICANT: ASADA, Kiyozo
22 ; APPLICANT: KATO, Ikunoshiin
23 ; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
24 ; FILE REFERENCE: TAKAKURA-6
25 ; CURRENT APPLICATION NUMBER: US/10/090,624
26 ; CURRENT FILING DATE: 2002-03-06
27 ; PRIOR APPLICATION NUMBER: 09/445,412
28 ; PRIOR FILING DATE: 1999-12-06
29 ; PRIOR APPLICATION NUMBER: 151969/1997
30 ; PRIOR FILING DATE: 1997-06-10
31 ; NUMBER OF SEQ ID NOS: 33
32 ; SOFTWARE: PatentIn version 3.0
33 ;
34 ; SEQ ID NO 4
35 ;
36 ; LENGTH: 522
37 ;
38 ; TYPE: PRT
39 ;
40 ; ORGANISM: Pyrococcus furiosus

```

```

Query Match 9.9%; Score 307.5; DB 10; Length 418;
Best Local Similarity 28.4%; Pred.No.3.1e-18;
Matches 116; Conservative 57; Mismatches 154; Indels 81; Gaps 20.

QY 139 SVEPLPYXIIPOIETFGKASLYKAXALDITQXNKREVOLEIEXIAQXXSNDYAYITA 198
      | : : : : : | : : : : : | : : : : : | : : : : : |
Db 49 SLSTFGSQDVAWMDGKBEFTTDDVAKQIQLQSNKMDIQKVENKMTVEVTEKAEVTA 108
      | : : : : : | : : : : : | : : : : : | : : : : : |
QY 199 KPEYK-----VMDVARGIKVADVAOSSYGLYGGGQIVAAVDVGLDTRGDDSSMEAF 251
      | : : : : : | : : : : : | : : : : : | : : : : : |

```



```
Db 109 VPSOTQPMGKISYND--QSITKT-----TGSGGIVAVLDYGVYTS-----HLLD 152
Qy 252 RGRITLVALGRTN-----NANDNGHGTIVAGSVLGNXTN---KMAPOANL-VFOS 301
Db 153 AGSAECCRFKPTOSNPLVDGSCDROGHGTHVACTVLAHSGSNGQGVYVAPQAKLMAIKV 212
Qy 302 IMDSXGGLG-GLPSNLQTLFQSAXSAGARITHN-SWGAIVNCAVTTDSRNVDDYVRKNDM 359
Db 213 LGNGSGYSDDLIAAIAIRHVADEASRTSKYVIMSGSSAKSLA---SAYDYAYGKV 269
Qy 360 TILFAGNEXPRNGTISAPGTAKNAITVGATEMLRPSGSYADNINHYAOSSRG-PTKD 418
Db 270 LIYAAAGNSGSSNTIGFPGGLVAVVAALLENVOON-GTY-----RYADFSKGNPATA 323
Qy 419 G-----RIPDYMAPGTYYILSARSSLAPDSSFWANHDSKYAYMGTSMAPTIVAGNYA-- 471
Db 324 GDYIIQERDIEVASPGASV-----ESTW--TYGGYINTISGTSMAPTIVAGLAAKI 371
Qy 472 -----QLREHFVKNRGITPKPSLLAKALIAGAAD--XGLGYP 506
Db 372 WSANTSLSHQLRTE-LQNRA--KYVDIKGIGAGTGDYVAGSGFYP 415

RESULT 13
US-10-090-624-6
; Sequence 6, Application US/10090624
; Patent No. US2002013235A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHTA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshi
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 1398
; TYPE: PRP
; ORGANISM: Pyrococcus furiosus
US-10-090-624-6

Query Match 9.7%; Score 299.5; DB 12; Length 1398;
Best Local Similarity 22.9%; Pred. No. 7 6e-17;
Matches 173; Conservative 76; Mismatches 248; Indels 257; Gaps 31;

Qy 96 LEETKQXLEXTGAKILD-----YIPDYATIVERE-----GDVXSXXXXIE 135
Db 77 LBAKTELEKLAELDENRNLNMLVKIRPEVKELNLTISLEKAMLNREVKLSPRIE 136
Qy 136 -HVESVEPLPYXIXIDPOLFT-----KGA SXLVKAXALDTOKNKEY 176
Db 137 KDKVTEP-----SLPPMNVSTWVINALQFIQEFYDSSGVVAVLDGTVDPMHPELSI 191
Qy 177 QLRGIEIXIAQXXXSNDVYITAKPEY-KVNN-----DVARGI----- 212
Db 192 TPDGRKRIIEMKDFDEGFEVDTSFSFSKVVNGTLIINTTFQVASGLTLNESTGLMEYVK 251
Qy 213 -----VKADVAQSSYGLY-----GGOGVVAAYDT 236
Db 252 TVYVSNVTIGNITSANGIYHGLLPERYFDLNFQDQEDFYVLLVNSTGCGYDIAVYDT 311
Qy 237 GLDTGKNDG-----SMHEAFRGKITALVALGTTNNAN-----DINGHGTIVAG 279
Db 312 DLQYDTDEVPYLGQYNTYVDAVAFSYTYGGLNLYVLAIEDIPNGEYAVFQMDHGHTYAG 371
```

```
Qy 280 SVLGNXGXTN-----KGAPOANLVPFOSIMDSXG 307
Db 372 TVAGYDSNNDADMDLMSYSGMEVEFSRLYGMDVTNTTDTVOGVAPOAIMAIRVLS-D 430
Qy 308 GAGGLPSNLQTLFQSAXSAGARITHNSWGAIVNCAVY--TOSRN--VDDYVRKNDMTIF 363
Db 431 GSGSMMDIIEGK-YYAANTHAGADYJMSLGG--NAPYLDGTBPESAVAADELREKYGVYFI 487
Qy 364 AAGNEPRNGGTISAPGTAKNAITVGATEMLRPSGSYADNINHYAOSSRG-PTKD 418
Db 488 AAGNEPRNGGTISAPGTAKNAITVGATEMLRPSGSYADNINHYAOSSRG-PTKD 418
Qy 406 HAAQSSRQPTKGRIRPDYMAPGTYYILSARSSLAPDSSFWANHDSKYAYMGTSMAPTI 465
Db 546 RIAPFSRQPTKGRIRPDYMAPGTYYILSARSSLAPDSSFWANHDSKYAYMGTSMAPTI 465
Qy 466 VAGNVAQLREHFVKNRGITPKPSLLAKALIAGA--DXGLGYPNGQMGVYTLQK 519
Db 596 VSGVYVALLISG-AKAEGLYYNPDIKKVLESGATWLEGGPYTGQRYTELDOGHLYNVTK 654
Qy 520 SLNVAVYVNESSXLSQKATYXFTATAGRPKLSLWSDAPASTTA--SVTLVNDLILV 576
Db 655 SWEI-----LKAINGTTLPIVDHMDKSYSDFAEYLGVDYIRGLYAR 696
Qy 577 ITAPN-----GTYVGN--DFXXPXXXXN-----DG-----RNVE----- 605
Db 697 NSIPDIEVNHIKYVDIEKRFELIATEPMIKPFVSGVILENTEFLRYKKYDVEGLER 756
Qy 606 -----NVFINPKQSGTYTIEVOAYNPVGPQXFS 634
Db 757 GLYVGRHIIIDPT--TPYIEDEILNTIVIPKFT 788

RESULT 14
US-09-927-827-59
; Sequence 59, Application US/09927827
; Publication No. US20030036176A1
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; APPLICANT: Ramseler, Thomas M.
; TITLE OF INVENTION: Directed Genetic Engineering of Xanthomonas campestris
; FILE REFERENCE: 38-10(15824)B
; CURRENT APPLICATION NUMBER: US/09/927,827
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/279,493
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 69
; SEQ ID NO 59
; LENGTH: 595
; TYPE: PRP
; ORGANISM: Xanthomonas campestris
US-09-927-827-59

Query Match 9.08%; Score 280.5; DB 9; Length 595;
Best Local Similarity 24.9%; Pred. No. 9.9e-16;
Matches 125; Conservative 64; Mismatches 163; Indels 151; Gaps 21;

Qy 92 NGPILSETKQXLEXTGAKILDYIPDYATIVEREGDVXSXXXXIEHVESVEPLPYXIXIDP 151
Db 136 NAVLTKIDLESEDLAENGA-----DYRHEVLEILD-----RRVIGAP--AGVA 176
Qy 152 QLETKGASLVKAXALDTOKNKEYVQLRG-----HEXIAQXXXSNDVYIT 197
Db 177 QLTTHNRQAGHSTGDLGDALDEAIELEDSLQRLISARLTRREIETLOSRYRALGFRRV 236
Qy 198 AKPEYKVNNDVARGIVKADVAQSSYGLYGGQIVAVADPTGIDTGRNDSMHEAF-----R 252
Db 237 RDAGKRALIHVSGVNLVNDAAARTAYRARGOQIGMAYLDGT-----AAHPHFVPPGR 290
Qy 253 GKITALLY-----ALGTNNANDTNGKGTIVAGSYLG-----NGXTNR- 289
Db 291 DNVAQMDCTRRGAPKRLTRADGKREFPALDGHGHTAGITAGCCRAAVPADSGKPGEL 350
```

QY 290 ----GMAPQANLVFQSIWDSXGLGLPS-----NLOTLEFSQA-----XSACAR 329
DB 351 LEFAGMAPETQLYGFKYLLDDAG--NGRDSWMKAKAVOHADINERAGELIVHGNLSLGGY 408
QY 330 IHTSMGCAAVGATTTDSRNVDDVRRKNDMTILFAAGNE-----XPNCGT-----I 375
DB 409 FDPESYCGGF-----TPLCINELRLRMROGLVLYVAAGNEGLAMLQONGGTYTPANMDLSI 463
QY 376 SAPGTAKNATVGCATENLRP-SFGSYADININHVAQSSRCPTKDRIKPDVMAFGTXILS 434
DB 464 SDPQNLEDAIVGVSVHNSPPHNYG-----VGYFSRGFTADGRSKPDVVAAGEKILS 515
QY 435 ARSSLADSSFWANHDSKYAYMGTSNATPIVAG-----NVAQL-----473
DB 516 AAYGFDP-----RDPSSIMVEMSGTSMAPHVSGVLAGFLSARREIFGPDPRVKQLLDT 570
QY 474 -----REHFVKNRGITTPKPSILK 491
DB 571 CTDLQRDRXYVGRGV---PNLMR 590

RESULT 15

US-09-927-827-55
; Sequence 55, Application US/09927827
; Publication No. US20030036176A1
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; APPLICANT: Ramseyer, Thomas M.
; TITLE OF INVENTION: Directed Genetic Engineering of Xanthomonas campestris
; FILE REFERENCE: 38-10(13824)B
; CURRENT APPLICATION NUMBER: US/09/927,827
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/279,493
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 69
; SEQ ID NO 55
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-09-927-827-55

Query Match 8.8%; Score 273.5; DB 9; Length 580;

Best local Similarity 22.7%; Pred. No. 3.8e-15;

Matches 159; Conservative 80; Mismatches 250; Indels 211; Gaps 30;

QY 1 MRKKKVLSTVSAALISTVALNPSAGXARXFDLDFKIQIOTTTDXGFSKXOTGAAP 60
DB 6 LRKRGTSLTILGASALTSLAMPAPAG---EYILDGLATPOTHOKEFIVYTKDGSTAL 60
QY 61 LVESENVKLXKGLAKKLETPVANKKLHXQFNGLIETKQXLEXTGAKIIDYIPDYVI 120
DB 61 ASFS---ALTSLKTAANAAPA-----KAGKALGINSYRLALGPELV-- 100
QY 121 VEYEGDVXSXXXIEHVESVEPYLPXYIDPOLFTGASXLVKAXALDTKXNKEVQLRG 180
DB 101 -----RADRALDRAEAEETLMROLAADPNVQSEV----- 129
QY 181 IEXIAQXXXSNDVXYITAKP-EYKVMNDVARGIYKA--DVQSSYGLGQGIYAVADTG 237
DB 130 -----DQILHATLTPMDTRLSEQMAFGTTNAGLINIRPAMDKATGSGVVAVIDTG 179
QY 238 L-----DTGRNDSMHEAFRGKITALYALGRTNAND-----TNGHGT 275
DB 180 ITSHADLANANTLAGYDFISDATTARDGNGRDSNMADEGDWYAANCGAGIPRASSSWHGT 239
QY 276 HVAGSV--LGNGXTN-KGMAPQANLVFQSIWDSXG-----LGGLPSNL 316
DB 240 HVAGTVAAVTNTTGVAGTAVGAKVVPVRYLGKGGSLDIADAIWVASGGTVSGIIPANA 299
QY 317 Q--TFEISOXSAGARIHNSMGAAVNGAVYTTDSRNVDDVYRKNDMTILFAAGNEXPNGCT 374
DB 300 NPAEYINSLGSGGSCSTTKQN-AINGAVSRGT-----TVVVAAGNDASNVSG 346

QY 375 ISAPGTAKNATVGCATEN--LRPSFGSYADININHVAQSSRCPTKDRIKPDVMAFGTXI 432
DB 347 -SLPANCANVIAVAATTSAGAKASYSNFGTGI-----DVSAPGSSI 386
QY 433 LSARSS--LAPDSSFWANHDSKYAYMGTSNATPIVAGNVAQLREHFVKNRGITPK--PS 488
DB 387 LSTLNSGTTTPGS-----ASVASYNGTSMASPHVAGVALVOS--VAPALTPAAVET 437
QY 489 LLK--AALLIAGNADXLGY-----PNGNGWGRVTLKSLNVAAYNESX 531
DB 438 LKKTARALPGACGGCGGAGIVNADAATVTAINGSGGGGGGNTLNGFTVGTGLG---- 493
QY 532 LSTOKATYXTATAGKPKLISLYWSDAPASTTASVTL---VNDLDLYT--TAPNGTXY 585
DB 494 AATGAELNVTITVPAG-----SGTLVTVTSGGSGDADLYVRAGSAPDTSAY 539
QY 586 VGNDPFXKXXXNMGRNVENVFINXQSGTYTIEVQAYN 625
DB 540 TCRPYRS-----GNAETCTTAP-SGTYVYVRLAYS 569

Search completed: April 1, 2003, 12:52:24
Job time : 26.9812 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:36:24 ; Search time 17.4863 Seconds
(without alignments)
3513.026 Million cell updates/sec

Title: US-09-920-954-1

Perfect score: 3101

Sequence: 1 MRKKVFELSLAAALSTFY.....EVOAVNPVGPQXSLATVN 639

Scoring table: BLOSUM62

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR_73:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	518.5	16.7	1743	2	multidrug resist
2	477.5	15.4	1905	2	intracellular alka
3	350.5	11.3	444	2	intracellular alka
4	319.5	10.3	442	2	pyrrolisin (EC 3.4.
5	297.5	9.6	1398	2	microbial serine p
6	286	9.2	806	2	subtilisin (EC 3.4
7	279.5	9.0	419	1	subtilisin-type pr
8	274.5	8.9	799	2	serine proteinase
9	273.5	8.8	420	1	cell wall-associat
10	268.5	8.7	580	2	surface layer-asso
11	268.5	8.4	715	2	high-alkaline seri
12	260.5	8.4	894	2	subtilisin (EC 3.4
13	257.5	8.3	1345	2	subtilisin (EC 3.4
14	255.5	8.2	378	2	subtilisin (EC 3.4
15	255.5	8.2	381	2	subtilisin (EC 3.4
16	253.5	8.2	513	1	subtilisin (EC 3.4
17	253.5	8.2	382	1	subtilisin (EC 3.4
18	253.5	8.2	381	1	subtilisin (EC 3.4
19	251.5	8.1	381	1	subtilisin (EC 3.4
20	250.5	8.1	380	2	high-alkaline seri
21	250.5	8.1	601	2	serine proteinase
22	250	8.1	382	2	subtilisin (EC 3.4
23	249.5	8.0	379	1	subtilisin (EC 3.4
24	249.5	8.0	757	2	subtilisin (EC 3.4
25	248	8.0	627	2	subtilisin-type pr
26	247.5	8.0	381	2	serine proteinase
27	246.5	7.9	402	2	subtilisin (EC 3.4
28	242	7.8	1331	2	probable surface 1
29	239.5	7.7	401	2	serine proteinase

30	239.5	7.7	519	1	alkaline proteinase
31	235	7.6	519	2	halolysin R4 (EC 3
32	232	7.5	488	2	protease (Import
33	231.5	7.5	1167	1	streptococcal C5a
34	230	7.4	384	2	alkaline proteinase
35	223.5	7.2	409	1	serine proteinase
36	223.5	7.2	321	1	alkaline proteinase
37	220.5	7.1	535	2	hypothetical prote
38	219.5	7.1	613	2	subtilisin (EC 3.4
39	218.5	7.0	272	2	subtilisin (EC 3.4
40	218.5	7.0	275	2	subtilisin (EC 3.4
41	218.5	7.0	533	1	serine proteinase
42	218	7.0	361	2	high-alkaline seri
43	216.5	7.0	361	2	halolysin (Importe
44	216.5	7.0	525	2	serine proteinase
45	216.5	7.0	645	1	SUBSMP

ALIGNMENTS

RESULT 1
T18279
multidrug resistance transport protein - slime mold (Dictyostellium discoideum)
C:Species: Dictyostellium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
R:Shaulsky, G.; Loomis, W.F.
Submitted to the EMBL Data Library, June 1996
A:Reference number: Z18855
A:Accession: T18279
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1743 <SHA>
A:Cross-references: EMBL:U60086; NID:g1399914; PID:g1399915; PIDN:AAB0331.1
C:Genetics:
A:Gene: tagC

Query Match 16.7% Score 518.5 DB 2: Length 1743;
Best Local Similarity 24.2% Pred. No. 2.3e-28;
Matches 194; Conservative 114; Mismatches 213; Indels 281; Gaps 33;

QY	87	HIQFNGPILEETKQXL-----EXTGAKILDPYDAVYIV--EYEGDYXN 130	
DB	115	YIVQFKRINDRETKLEFLIGTDIVDEOPYOSHIVHYIPHDSFLVMTQEOSVYLS 174	
QY	131	XXXIEHVESVEPY-----LPYXIDPOLFTKGASXLVK--AKALDT--KQXN 173	
DB	175	KENVSMIGFEPSNKHILVNEKSIQLPVYLI---LDSNTSLIQREMENTLSILKSV 230	
QY	174	KEYOLGI-----EXINQXXSNDVXYXTAKPEVK 203	
DB	231	SKYKLLNQKKLAKSYVCNDESPSPCSLINESEKLYQWISQSESN--FIRSEKQ 287	
QY	204	VMDVARGIV-----KADVAQSSYGLYGQQLYAVADTGLDTGR---NDS---- 245	
DB	288	TARLSRPKVFSTKPTLVNDRDIP---LKGKQILSLINDTGLDSCFESDSTAS 342	
QY	246	---SMHEAFRGKITVALVGLFTNNANDTNGCHTHVAGSVLG-----NGXTNKGMAPOA 295	
DB	343	PLNSVNLNR-KVVTYITSTSDSDKVDHGHILCGSAAGPREDSSVNTSSFFGLATPA 401	
QY	296	NLVFQSIKMSXGLGL--PSNLOTFLSQAXSAGARIHNSWGA---AVNGATYDSRN 349	
DB	402	KIHF---FDLASGSSSLTPPDSKLTQPLYDAGARVHDSGVSVEGYGYSPTAS 458	
QY	350	VDDYVARN-DMTILFAAGNEXPCGTIS--APGTAKNATVGAFTENLR-----SPFS 399	
DB	459	IDQFLTHDPFILRAAGN---NEQYISLTQSTAKNVITVGAHOTIHEVNDLDPNVI 515	
QY	400	YANI----- 404	
DB	516	YQSSVDINQELCDPDSRYCNYTTAQCCLSNATGCLASCCPTLLKRSVIDANNTQPLLY 575	

OY	405	--NHVAQSSSKOPTDGIKEDVNAPEXTIILSARSLA-----PDSSFMANHDSKAYVM	456
Dd	576	NENNICSSSKSPPTIDHGEMKPALVAPGEYITSARSNGANTTDOCGDSL-ENTNALLA-I	633
OY	457	GGETSMATPIVAGNVAQLREH-----FYKNGLIPKPSLLKAALLAGA-----	498
Dd	634	SOTSMATSEPAATAATTILKOYLVDGYPPGTSIVESKKLOPOTSLALMINNALINGTFQ	693
OY	499	--ADGLGYPNGN-----OGMGRVTLDKSLNAAVNESS-----	530
Dd	694	LITSSIIYPPSQOVEENFAGASLYQGMCALIRSNMLHYNNNNNSNNNKTKSDGITKEPCI	753
OY	531	-----XUSTOKATYAFYFT-----ATAGRPKR---ISLWSDAPASTT	564
Dd	754	GGLDLRLVKPNQMKESISTGONTSYCFYKPKSSSSSGNNIIPRYAVATLVWTDPSPVAG	813
OY	565	ASVTLVNDLDLYI-----TAPNGIXYYGNDFXXPXXMXWDGRNVENVFIIXP	612
Dd	814	AKFNLYNNLDLTMIYYRDNGSTIFYSSNOGGSBFLG-----LAPQDTLNINVEGIVHNPT	867
OY	613	OSGTYTIEVQAANVPVGPOXES	634
Dd	868	EPMTRYREMVAQTNPVGPONES	889

RESULT 2

multidrug resistance protein - slime mold (*Dictyostellium discoideum*)
C:Species: *Dictyostellium discoideum*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18267
R:Shaalsky, G.; Kuspa, A.; Loomis, W.F.
submitted to the EMBL Data Library, January 1995
A:Description: An MDR transporter/serine protease gene is required for prestalk speciality
A:Reference numbers: Z18850
A:Accession: T18267
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1905 <SWA>
A:Cross-references: EMBL:U02432; NID:g664839; PID:g664840; PIDN:AAA62212.1
C:Genetics:
A:Gene: tagp

Query Match	Score	DB 2	Length
Best Local Similarity	15.48	477.5	1905
Matches 193, Conservative	24.58	Pred. No. 2.1e-25	
	98	Mismatches 237	Indels 259; Gaps 31;

```

Oy  87 HIXORNGILBETKXLE-----XTCAKLIDYIPYAYIYER-----EGDYXS 129
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  172 YIVQKRIINBETROKLEFLIGLDTITLKBQPKSHIVHPIPHDSFLVMTKEQSVLLS 231
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy  130 XXXXIEHESVEPY-----LPXYXI-----DPOLETKASXL-----V 162
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  232 SKEMTSMIGEHFPSKIKILNTHKSICTPYVYIILSGTNSLIQREWNLTNSLITSYNSKV 291
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy  163 KAXALDTKO-----XNKEVOLRGIEIXIAQXXXS--NDVXYITAKPEYKVM 205
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  292 KUTLINOQKLKSIYVCNDESPSSSSSSCSLIGSKIKIYKMIISBQSESENIYERSKLOTA 351
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy  206 NDVANGIV-----KADVAOSTYGLYGOGIYAVADTGLDTGR--NDS----- 245
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  352 NKLSPFYFETGRDKLVNNDRIDIP-----LRGKGQILSIADTGLDGSCHFPSDSKYPIF 406
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy  246 -SMHEAFGKITALYALGORTNANNANDNTNGHGTNVAGSVLGNKXTN-----KGAAPQANL 297
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  407 NOVNNENHKKVYT---YIITYHDNEDYVNGHGTNVGCSAAGTPEBDSWMAISSFSGLATAKI 463
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy  298 VFQSTIMDSXGIGLGPNSLIQTLFESQAXSAGARIHNSNGA-----AVNCAVYITDSRNVDDY 353
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  464 AFYD-LSSGSSSPPTPEYISQMYKPLIYAGARVHSDMSGVSISLQGYIGYIGSDDGAGIDAF 522
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy  354 VRK-NDMTILFPAAGNEXDNGGTTISAPGTAKNAITVGATENTLRPSFGS-----YADNI- 404
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

[illegible]

RESULT 3

intracellular alkaline serine proteinase aprX [imported] - *Bacillus halodurans* (strain C:Species: *Bacillus halodurans*
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: B83891
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a
A:Reference numbers: A83650; MUID:20512582; PMID:11058132
A:Accession: B83891
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <STO>
A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA05649.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: aprX

Query Match	11.3%	Score 350.5;	DB 2;	Length 444;
Best Local Similarity	30.28;	Pred. No. 2.8e-17;		
Matches 114; Conservative	54;	Mismatches 127;	Indels 83;	Gaps 17;

```

0Y 162 EXIAOXXXXXNDVXYITAKPEYKVMNDVARGLYKA-DVAOSSYGLXGQGYAANDPGLDIT 240
Db 100 ESIQOMLVCKCQIDIRKLYLRENVHALDITAVESAQAPEYIRNETTITGADVTITAVITDIT 157
0Y 241 GRNDSMEAEFRGKITALLY-ALGRTNANPDNTNGHTVAGSVLNGXTN---KGMAPOA 295
Db 158 ----YPHEDELEGRIKAFVDEVNOREEPYDONGHTHCAGDMAANGASSDPOYRGAPAEA 212
0Y 296 NLVFGSINDSXGGLGGEPSNLQTLFSGQAXSAGARLHTNSWCAAVNGAYTTPDSRVV----- 350
Db 213 NVIGKVKVILNKQ-GMOSLESIIMQV-----EMCIQYNEERHPDDPHIISMSL 257
0Y 351 -----DDIVR-----KNDKTILPAGNEXPNGCTISAPGKTAKNATVGTEN 392
Db 258 GGOALPYENEGEDDPVYRILVEEAMNAGITVCVAAAGNSGDAOTIASPGVSEKIVITVGLALDD 317
0Y 353 LRFSGSYADININHVAOFSSRGPTKDGRIKPDVNAFGTXILISARSSLAPDSSF-----WA 447

```

Db 318 -RDTTDEDD---VAPSSRGPTTYGKPKFDLLAPGVNIVSLRS---PNSFYDKIQKGS 370
Oy 448 NHDSKYAYMGSTMATPIVAGNVNQLREHFVKNKGTIPKPSLKAALLAGA---AD----- 500
Db 371 RVGSHHTYMSGTSMTAPVACGAVVALLMLQH---EPNLTPOE--VKTRLMGCTDRMADRPDN 425
Oy 501 -XGLGY-----PNCNQ 510
Db 426 VYGAGYISAGCAIPNSEE 443

RESULT 4

Intracellular alkaline serine proteinase aprx - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999

C:Accession: A69587

R:Kunst, P.; Ogawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emerson, P.T.; Enlian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Gallizi, A.; Gall

lechi, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koether, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadachi, J.; Sato, T.; Scanlon

A:Authors: Schleich, S.; Schreuter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, K.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wilpat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zunshtein, E.; Yoshikawa, H.; Danclun, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: A69587

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1442 <KUN>

A:Cross-references: GB:299113; GB:AL009126; NID:g2634090; PIDN:CAM13610.1; PID:el183385;

A:Experimental source: strain 168

C:Genetics:

A:Gene: aprX

C:Superfamily: subtilisin homology

F:146-398/Domain: subtilisin homology <SBT>

Query Match 10.3%; Score 319.5; DB 2; Length 442;

Best Local Similarity 30.0%; Pred. No. 4.4e-15;

Matches 106; Conservative 46; Mismatches 108; Indels 93; Gaps 15;

Oy 201 EYKMNDAVAGIYKA-DVAOSSYGLYGQGIYAVADTGLDTGRNDSMHEAFRGKITALY 259
Db 120 EVALLDITATEASHAKEVVRNQGTLTGKGVYAVVDTGI-----YHPDLEGRI----- 168
Oy 260 ALGRTNNAN-----DTNGHGHVAGSVLNGXTN-----KGNAPQANLVFQSIIMDSXGL 309
Db 169 -IGFADVVKQKTEPYDDNGHGHGAGDYASGASSGQYGRAPENLGVVLLKQGS- 226
Oy 310 GGLPSNLQTLFSGAXSAGARITHNSMGAAVN-----GAYTDSRANVD 352
Db 227 GTLADIIEGV-----EMCIQYNEDEPDEPIDIMSGIGALALYDHEQDP 272
Oy 353 YVKKND-----XTILFAAGNEXPNGTISAPGTAKAITYGTEMLRPSFGSTADNINH 406
Db 273 LVRAVEASAGIVCAVAGNSGPDQTIASPGSEKVIYVGLADNNTA-----SSDDT 328
Oy 407 VAOFSRGPTKDGRIKPDVMAPIYLLSARSLPADSSF-----WANHDSKYAYMGSTM 461
Db 329 VASFSHSGPVTYGKERPDILAGVNIISLRS---PNSYIDKLOKSSRSVSOQFTMSGTSM 385
Oy 462 ATPVAGNVAQLREHFVKNKGTIPK--PSLKAALLAGAAGXGLGPNQNM 512
Db 386 ATPICAGIALLIQ---QNPDLTPEVKEVLLK-----NCTDKM 420

RESULT 5

128159

Pyrolysine (EC 3.4.-.-) - Pyrococcus furiosus

C:Species: Pyrococcus furiosus

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 24-Oct-2000

C:Accession: T28159

R:Voornhorst, W.G.B.; Eggen, R.J.L.; Geerling, A.C.M.; Platteuw, C.; Slezén, R.J.; Vo

J.; Biol. Chem. 271, 20426-20431, 1996

A:Title: Isolation and characterization of the hyperthermostable serine protease, pyr

A:Reference number: 220481; MUID:96355370; PMID:8702780

A:Accession: T28159

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 11398 <VOO>

A:Cross-references: EMBL:U55835; NID:g1556462; PID:g1556463; PIDN:AAB09761.1

A:Experimental source: DSM3638

C:Genetics:

A:Gene: pls

C:Keywords: hydrolase; serine proteinase

Query Match 9.6%; Score 297.5; DB 2; Length 1398;

Best Local Similarity 22.9%; Pred. No. 7.6e-13;

Matches 173; Conservative 76; Mismatches 248; Indels 257; Gaps 31;

Oy 96 LEETKQXLEXTAKILD-----YIPDAYIVEE-----GDYXSSXXXIE 135
Db 77 LEEKTELEKLAELIDENRVLMVLKIKPEKVKELNYSLEKAMLNREKXLSPIVE 136
Oy 136 -HVESVPEPLPYXIDPOLFT-----KGASXLVKAAXALDTQXKKEV 176
Db 137 KDVKTREP-----SLEPKMYSTWVINALQIQIEFGYDSSGVVAVLDTCVDPNHPFLSI 191
Oy 177 QLGIEKIXQXXNSNDYVYITAKREV-KVMN-----DVARGI----- 212
Db 192 TPBGRRKILIMKDFDEGFVDTSEFSKYVANGILINTTPQVASGLITLNESTGLMEYVK 251
Oy 213 -----VKADVAOSSGLY-----GQGIYAVADT 236
Db 252 TVYVSNTYIGNTISANGIYHFGLLPERYFDLNFQDQEDPYVLLVNSTGNGYDIAVDT 311
Oy 237 GLDTGRNDS-----SMHEAFRGKITALYALGRTNNAN-----DTNGHGHVAG 279
Db 312 DLQYDPTDEVPLQGYNTYDYVAVESYGGPLNYVLAIEDPNEGEYAVFGMDGHGHTVAG 371
Oy 280 SVLNGXTN-----KGNAPQANLVFQSIIMDSXG 307
Db 372 TVAGYSDNDAMDMLSMYSGMEVEFSRLYGMWDTYNTVTDVQVAGAOIMAIRVLS-D 430
Oy 308 GLGGLPSNLQTLFSGAXSAGARITHNSMGAAVNGAYT--TDSRN--VDDYVRKNDMTILF 363
Db 431 GRGSMMDIIEGM-TYAAVHGADVISMSLGG--NAPYLDGTDPESAVDELTEKYGAVFI 487
Oy 364 AAGNEXPNGTISAPGTAKAITYGATEMLRPSFGSYAD-----NIN 405
Db 488 AAGNEXPNGTISAPGTAKAITYGATEMLRPSFGSYAD-----NIN 405
Oy 406 HVAQFSRGPTKDGRIKPDVMAPIYLLSARSLPADSSPMANHDSKYAYMGSTMATPI 465
Db 406 HVAQFSRGPTKDGRIKPDVMAPIYLLSARSLPADSSPMANHDSKYAYMGSTMATPI 465
Oy 546 RIAFSSRGPRIDGELKRPVAVAPGIGYSLPMKIGAD-----MSGSMATPH 595
Db 546 VAGNVQALREHFVKNKGTIPKPSLKAALLAGA-----DXGLGYPNQNM 512
Oy 520 SLNVAAYVNESSXLSTQKATYXFTATAGKPLKISLVMSDAPASTTA---SVTLVNDLVL 576
Db 520 SLNVAAYVNESSXLSTQKATYXFTATAGKPLKISLVMSDAPASTTA---SVTLVNDLVL 576
Oy 576 ITAPN-----GTYYVGN-----DFXXPXXXNN-----DG-----NNNVE----- 605
Db 576 ITAPN-----GTYYVGN-----DFXXPXXXNN-----DG-----NNNVE----- 605
Oy 697 NSIPDLIVEWHIKYVGDTEYRTFEIYATEPMIKPFVGSVILNNTEFVLRVKYDVEGLEP 756
Db 697 NSIPDLIVEWHIKYVGDTEYRTFEIYATEPMIKPFVGSVILNNTEFVLRVKYDVEGLEP 756
Oy 606 -----NVFINXPOSGTYTIEVQAVNPVGPQXFS 634

Db 757 GLYVGRITIDPT--TPYIEDEILNTIPIPEKFT 788

RESULT 6

A41341

microbial serine proteinase (EC 3.4.21.-), minor (vpr), precursor - Bacillus subtilis

C/Date: 17-Jul-1992 #sequence.revision 17-Jul-1992 #text_change 20-Jun-2000

C/Accession: A41341: B41341: S39700: D69730

R:Stroma, A.; Rufo Jr, G.A.; Theriault, K.A.; Dwyer, M.; Wilson, S.W.; Petro, J.

J. Bacteriol. 173, 6889-6895, 1991

A>Title: Cloning and characterization of the gene for an additional extracellular serine

A/Reference number: A41341; MUID:92041574; PMID:1938892

A/Accession: A41341

A/Molecule type: DNA

A/Residues: 1-806 <SID>

A/Cross-references: GB:M76590; NID:g143819; PIDN:AAA22881.1; PID:g143820

A/Accession: B41341

A/Molecule type: protein

A/Residues: 161-195 <SL2>

R:Glasier, P.; Kunst, F.; Arnaut, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu,

, A.; Rapoport, G.; Danchin, A.

Mol. Microbiol. 10, 371-384, 1993

A>Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fr

A/Reference number: S39655; MUID:95020537; PMID:7934828

A/Accession: S39700

A/Molecule type: DNA

A/Residues: 1-806 <GLA>

A/Cross-references: EMBL:X73124; NID:g413923; PIDN:CAA51601.1; PID:g580871

R:Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Buschli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

lech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lader, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maveel

Y, M.; Ogawa, K.; Ogatawara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Potteille

Rieger, M.; Rivolta, C.; Roche, B.; Roche, M.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A:Authors: Schlecht, S.; Schroter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: D69730

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-806 <KUN>

A/Cross-references: GB:299123; GB:AL009126; NID:g2636240; PIDN:CAH15835.1; PID:g2636344

A/Experimental source: strain 168

C/Comment: The amino terminal sequence of the mature protein and a molecular weight of C

C/Genetics:

A:Gene: vpr

A:Start codon: TTG

C/Superfamily: microbial serine proteinase vpr; subtilisin homology

C/Keywords: hydrolase; serine proteinase

F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-160/Domain: propeptide #status predicted <PRO>

F:160-548/Domain: subtilisin homology #status atypical <SPT>

Query Match

Best Local Similarity 9.2%; Score 286; DB 2; Length 806;

Matches 132; Conservative 60; Mismatches 186; Indels 206; Gaps 19;

Db 145 MKDKVITSEDAVSPQMDASPIYGANDAWMDLGYTGKIKVAILDITGVEYNHPDLKKNFG 204

Qy 243 -----NDSSMHEAFRGKITALVALGRTNANDNGHCTHYAGVGLNGXTKGMAR 293

Db 205 QYKGYDFVNDNDYDKERTPTG-----DPRGEATDHTGTHVAGTVAANG--TIKGVAP 252

Qy 294 QANLVFOSIMDSXGGLGIPSNLOTLFSQASACARLHTNSWGAAYGATTTDSRNDY 353

Db 253 DATLLAVRVGLPGG--SGTTEENVLAGVERAAVDADVNNLSLGNLNNPDMATSTAL-DW 309

Qy 354 VRKNDMTLFAAGNEXPNGTISAPGTAKNAITVGATE---NLRPSFGSY----- 400

Db 310 AMSGCVAAVTSNGNSGRNGTIVSGPTSRRLAISGATQPLNEVAVYFGSSSAKWGYN 369

Qy 401 -----ADNI----- 404

Db 370 KEDDVKALNKEVELVEAGIGEAQDFEGKDLTKVAAVVRKGSIAFVVDKADNAKKAIGM 429

Qy 405 -----NHVAQF 410

Db 430 VVYNNLSGEIEANVPNGMSVPTIKLSLEDGEKLVSAKLAGETTKFTLVSKALGEQYADF 489

Qy 411 SSRGPTKD-GRKPDVMAPGTXILSARSLAPDSSFMANHDSKYAYMGTSMATPIVAGN 469

Db 490 SSRGPDMDTMKPDISAPGVNIYSTIPTHDP-----HYGYSKQGTSMASPHIAGA 543

Qy 470 VAOLREHEVKNRGITPKPSL--LKAALTAGAA---DXGLGYPNGOGMGWRTLDKSLNV 523

Db 544 VAVIKQ-----AKPKMSVQIQKATIMNTAVTLKDSGGEVYPHNAQAGASARI--MNA 593

Qy 524 AYVNESSXLSSTQKATYFTFTACKPLKISLVSDAPASTASTYL 569

Db 594 --IKADSLVSPGSYSYGFLEKENGNETKNETFTIENOSIRKSYTL 637

RESULT 7

S25835

subtilisin (EC 3.4.21.62) precursor - Bacillus sp. (strain TA41)

C/Species: Bacillus sp.

C/Date: 22-Nov-1993 #sequence.revision 20-Feb-1995 #text_change 18-Jun-1999

A/Accession: S25835

R:David, S.; Feller, G.; Narinx, E.; Gerday, C.

Gene 119, 143-144, 1992

A>Title: Sequence of the subtilisin-encoding gene from an antarctic psychrotroph Bact

A/Reference number: S25835; MUID:93012966; PMID:1398082

A/Accession: S25835

A/Molecule type: DNA

A/Residues: 1-419 <DAV>

A/Cross-references: EMBL:X63533; NID:g40198; PIDN:CAA45096.1; PID:g40199

C/Superfamily: subtilisin; subtilisin homology

C/Keywords: extracellular protein; hydrolase; serine proteinase

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-110/Domain: propeptide #status predicted <PRO>

F:111-419/Product: microbial serine proteinase #status predicted <SPT>

F:135-373/Domain: subtilisin homology <SPT>

F:144,184,359/Active site: Asp, His, Ser #status predicted

Query Match

Best Local Similarity 9.0%; Score 279.5; DB 1; Length 419;

Matches 97; Conservative 47; Mismatches 123; Indels 61; Gaps 16;

Qy 171 QXNKEVOLRGITEXIAQXXXSDVXYITAKPEKYANDVA-----RGIVKADVAQSS 221

Db 81 QNNKNTLVEKPEL-----EIAATATKPE-ALYNMAAASQTPWGIKATYNNNSLST 132

Qy 222 XGIVGQGOIVAVPDTGDTGRNDSMHE-FAVRGKITLALYALGRT--NNANDTNGHGVH 277

Db 133 SG--GAGINIVAVLDGVTNHPDLSSNNVEQCKD-----FTVGNFTDNSCTDRGCHGHV 185

Qy 278 AGSVLNGXGTNKK--GNAPQANLVFOSIM--DSXGGLGELPSNLOTLFSQASAGARLHT 332

Db 186 AGSALANGCTSGSYGVAPREADLMAVYKVLGDDGSGYADDIDEALRHAGDQATALTNTKVI 245

Qy 333 N-SWGAAYNGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAITVGATE 391

Db 246 NMSLGSSESSLIT---NAVDYADKCVLLIHAAGNSGRKPSISIPYALVNAVAVALE 302
QY 392 NLRPSFGYADININVAOFFSRGPTKDG-----RIKPDYMAPCYXILSARSLAPDSF 445
Db 303 NTION-GTY-----RVADFSRGRKRTAGDYVIOKGDVEISAPGAAYST----- 346
QY 446 MANHDSKYAYMGCTSMATPIVAGNVAOL 473
Db 347 W-FDGGVATISGTSMASPHAAGLAKI 372

RESULT 8

subtilisin-type proteinase (EC 3.4.21.-) vpr precursor [similarity] - Bacillus halodurans
C:Species: Bacillus halodurans
C>Date: 04-Dec-1992 #sequence, revision 01-Dec-2000 #text, change 15-Jun-2001
C/Accession: G83753
R:Kunitz, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; M0ID:20512582; PMID:11058132
A/Accession: G83753
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-799 <STO>
A/Cross-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA04550.1; GSPDB:GNOC
A/Experimental source: strain C-125
C/Genetics:
A:Gene: vpr
C:Superfamily: microbial serine proteinase vpr; subtilisin homology
C/Keywords: hydrolase; serine proteinase
F:1-29/Domain: signal sequence #status predicted <SIG>

Query Match 8.9%; Score 276.5; DB 2; Length 799;
Best Local Similarity 23.5%; Pred. No. 1,1e-11;

Matches 199; Conservative 61; Mismatches 199; Indels 227; Gaps 25;

QY 77 LEIVPANKLHIXPNC-ILE-----ETKQLEXTGAKILDYIPDYIYEYE 124
Db 50 LEIVPANKLHIXPNC-ILE-----ETKQLEXTGAKILDYIPDYIYEYE 100
QY 125 GDVXKXKXIEHVESVEPYLPXYIDPQLEFTKASXLVAKALDTRKXNKEVOLRGTEI 184
Db 101 YDLFSGFALF-----LPAHQIPSLIGIDGNAHVYPIEYETDFEDV----- 144
QY 185 AQXKXNDYVYITAKPEYK-VYNDVARGIVKADVAOSSYGLGQGIYAADVGLDTRGN 243
Db 145 -----VIEKDAYSPMLDSAPFICANDAMEAGY--TGEIIVAILDITGVDTYHP 191
QY 244 DSSMHEAFRGKITALYAGRTNNANDT-----NGHGTAVAGSYLNGXTKKGAPQ 294
Db 192 D-LVNAF-GDYKGMDFIDNNDDPQETPPDPRGIEITHTGTHVACTVANAAGLI-KGYAPD 247
QY 295 ANLYPQSIINDSGIGLIPSLQTL--FSQAKSAGARITITNSGAAVAVGATTTDSRVND 352
Db 248 ANLALRYL-----GPGGSGTAGYVAGIERAVODGAGNNLSGLTLDLPATSTAL-D 302
QY 353 YAKNDMTILFAGNEKXNGGTISAPGTAKNAITVAGATENLRPSFGSY-----ADNINH 406
Db 303 WAAABGVAVVTSNGNSGNNMTVGSPTGSHDAISGAT--PLPYNKYKASVPTSDIDY 359
QY 407 ----- 406
Db 360 PSADIWGPPSDELLDELGETEYFAFAGLGKPGDFGVDEGKIALIVGEIPEVEKAEN 419
QY 407 ----- 406
Db 420 AKAAAGVAGIIVNNAVGVQPTVPGIAIPTIMLSNEDGLKMRNLENGONTVTFSEFDKL 479
QY 407 -----VQFSSRGPT-KDGRIPKPDYMAFGTYILSARSLAPDSFMANHDSKYAVMGST 461
Db 480 VGETVADFSRGRKRTAGDYVIOKGDVEISAPGAAYST-----GYGSRGQISM 533

QY 462 APPYVAGNVAOLREHFVKNRGITPKPSLLKALLAGAA-----DXGLGYPNGOGWRVTL 517
Db 534 ASHVAGAAALLLEAH-PNMGV-----DHVKAALMTAEVLVDENGRYPHNTOGAG----- 584
QY 518 DKLNAVAVNESSXSTOKATY-XFATAGKPLK 551
Db 585 --SIRIVAISETLVTPGSHSFGTTFKRGKQVE 617

RESULT 9

subtilisin (EC 3.4.21.62) 1 precursor - Bacillus sp. (strain TA39)
S23407
C:Species: Bacillus sp.
C/Date: 04-Dec-1992 #sequence, revision 04-Dec-1992 #text, change 18-Jun-1999
C/Accession: S23407
R:Marlin, E.; Davall, S.; Feller, G.; Gerday, C.
Biochim. Biophys. Acta 1131, 111-113, 1992
A>Title: Nucleotide and derived amino acid sequence of the subtilisin from the antarctic
A/Reference number: S23407; M0ID:92256481; PMID:1581352
A/Accession: S23407
A/Molecule type: DNA
A/Residues: 1-420 <NAR>
A/Cross-references: EMBL:X62369; NID:g40200; PIDN:CA44227.1; PID:g40201
C/Genetics:
A:Gene: subtil
C:Superfamily: subtilisin; subtilisin homology
C/Keywords: extracellular protease; hydrolase; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-111/Domain: propeptide #status predicted <PRO>
F:112-420/Product: microbial serine proteinase #status predicted <KAT>
F:116-374/Domain: subtilisin homology <SBT>
F:145,185,360/Active site: Asp, His, Ser #status predicted

Query Match 8.9%; Score 274.5; DB 1; Length 420;
Best Local Similarity 28.3%; Pred. No. 6.3e-12;

Matches 107; Conservative 56; Mismatches 142; Indels 73; Gaps 21;

QY 167 LPTKQXN-KEYQLRGIEIXIAQXKXNDYVYITAKE-YKVM--NDVARGIV---KA 215
Db 73 MEKQFNALKRKNLTYEKPEL---EATATDREALYNMAASQTPWGIKATYNNNS 128
QY 216 DVAOSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITAL--YALGRT--NNANDT 270
Db 129 STQRTS---GGGGINIVAVLDITGVNTN-----HPDLRNVECKDFVTGTYTNNSCDTR 179
QY 271 NGHGTAVAGSYLNGXTKXNK--GMAPOANLVFOSIK--DSXGLGLGIPSLQTLFSGAXS 325
Db 180 QGHGTHVAGSALADSGTGNCGYGAAPDADLMAYKVLGDGSGYADTIAAALIRHAGQDATA 239
QY 326 AGARITHTN-SMGAAVNGAYTTDSRVNDVYRKNDMTILFAAGNEKXNGGTISAPGTAKNA 384
Db 240 LNTKYIINSLSSSESLITNAVN--YSYKGVILITAAKNSGPTQSGISIGIPGALVNA 296
QY 385 ITVGATENLRPSFGSYADININVAOFFSSRGPT-KDG-----RIKPDYMAPCYXILSARSS 438
Db 297 VAAVALEN-KVENGY-----RVADFSRGRKRTAGDYVIOKGDVEISAPGAAYST--- 347
QY 439 LAPDSSFMANHDSKYAYMGCTSMATPIVAGNVAOLREHFVKNRGITPKPSL-----L 490
Db 348 -----W-FDGGVATISGTSMASPHAAGLAKIWAQYPSASVNDVAGELQYRAVENDI 398
QY 491 KAALTAGAAD--XGLGY 505
Db 399 LSGYYAGYGDPPASGQGF 416

RESULT 10

S11890
serine proteinase (EC 3.4.21.-) precursor, extracellular - Xanthomonas campestris pv.
N:Alternate names: subtilisin-related protease
C:Species: Xanthomonas campestris pv. campestris
C/Date: 21-Nov-1993 #sequence, revision 07-Feb-1997 #text, change 03-Dec-1999

C:Accession: S11890
 C:Date: Y.N.: Tanq, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.
 Mol. Gen. Genet. 220, 433-440, 1990
 A:Title: A multipurpose broad host range cloning vector and its use to characterise an
 A:Reference number: S11890; MUID:90251253; PMID:2187155
 A:Accession: S11890
 A:Molecule type: DNA
 A:Residues: 1-580 <LNU>
 A:Cross-references: EMBL:X51635; NID:g48533; PIDN:CAA35962.1; PID:g48534
 A:Experimental source: Xanthomonas campestris pv. campestris
 A:Note: The sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-Alt
 C:Superfamily: subtilisin; subtilisin homology
 C:Keywords: extracellular protein; hydrolase; serine proteinase
 F:1-32/Domain: signal sequence #status predicted <SIG>
 F:168-423/Domain: subtilisin homology <SBT>

Query Match 8.8%; Score 273.5; DB 2; Length 560;
 Best Local Similarity 22.7%; Pred. No. 1.2e-11;
 Matches 159; Conservative 80; Mismatches 250; Indels 211; Gaps 30;

QY 1 MRKKVFSLVSAALISTVALNPNAGXARXPDLPFGTGTDTTDXGFSKQXGTGAAF 60
 Db 6 LKRTGSLTIGASALTSLAMPAG---EYVDGLTAQTQKQFIVTYADGDTAL 60
 QY 61 LVESENKXKXGLKXKLETPVANKLHXPNPFLLETKOXLEXTGAKILDYIPDYAI 120
 Db 61 ASPS---ALTTSLRTAARAVPA-----KAKALGLNSVRRLALGPELV-- 100
 QY 121 VERGDVXAXXXIENHESVEPIYPIYXIDPQLFTKASXLYKXALDTPKXNKEVOLRG 180
 Db 101 -----RADRALDRAEAEETLMROLADPNVQSEV----- 129
 QY 181 IEIXAQQXXSNDVXYITAKP-EYKVMNDVARGIYKA--DVQSSYGLGOGQIYVAVDG 237
 Db 130 -----DQILHATLPDTRLSQWAFGTTNAGINIRPADKAGSSITVAVIDTG 179
 QY 238 L---DTGRNDSMHEAFRGKITALYALGRTNAND-----TNGHGT 275
 Db 180 ITSHADINANILAGYDFISDATPTARDGGRDSDNADEGDWYAANECAGIPAASSSMHGT 239
 QY 276 HVASV---LGGXKTN-KGMAQANLVFQSIIDMSXG-----LGLPBNL 316
 Db 240 HVAQTVAALVNTTGVAGTAIGAKVVPVRYLGKCGSLSDIADAIWASGGTVSGIIPNA 299
 QY 317 Q--TLFSAQASAGARIHTNSGAAVNGAYTDSRNVDDYKRNMTILFAENEXPNCT 374
 Db 300 NPAYVIMNSLGGGSGCSTTQMN-AINGAVSRGT-----TVYAAGNDASNVSG 346
 QY 375 ISAGTAKNAITVGATEN--LRPSFGSYADNINHVAQESSRGPTRKDGRIKPDVNAPTXI 432
 Db 347 -SLPANCANVAIVAAITTSAGAKASYSNCTGI-----DVSAFGSSI 386
 QY 433 LSARSS--LAPDSSFMANHDSKYAYMGSTMAPTPIVAGNVOLREHPFKNGITPK--PS 488
 Db 387 LSTINSGTTPGS-----ASIASYNGTSMASPHAVAGVALVOS--VAAPALPAAVET 437
 QY 489 LK--AALIAAGADXXGLGY-----PNCNOGWRVTLDSKLNVAAVNESSX 531
 Db 438 LKMTARLPACSGGGGAGIVNADAAVTALINGSGGGGGMTLTNGTPTVTGLG---- 493
 QY 532 LSTQKATYATATAGAPKLSLWSDAPASTTASVTL--VNDLDVY--TAPNGTX 585
 Db 494 AATCAELNVTITVPAG-----SGTLVTTSGGSGCDADLYVRAGSAPTDSAY 539
 QY 586 VGNDFFXPPXXNMGRNNVENVFINPQSGTYIEVOAYN 625
 Db 540 TCRPIRS-----GNAETCTITAP-SGTYIVRLKAYS 569

RESULT 11
 JCA4908
 alkaline serine proteinase (EC 3.4.-.-) I precursor - Alteromonas sp.
 N:Alternate names: sulliase

C:Species: Alteromonas sp.
 C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 08-Oct-1999
 C:Accession: JCA4908
 R:Tsujibo, H.; Miyamoto, K.; Tanaka, K.; Kaidzu, Y.; Imada, C.; Okami, Y.; Inamori, Y
 Biosci. Biotechnol. Biochem. 60, 1284-1288, 1996
 A:Title: Cloning and sequence analysis of a protease-encoding gene from the marine ba
 A:Reference number: JCA4908; MUID:97141200; PMID:8987544
 A:Accession: JCA4908
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-715 <TSU>
 A:Cross-references: DDBJ:D38600; NID:g1536787; PIDN:BAI18912.1; PID:d1019647; PID:g21
 A:Experimental source: strain O-7
 C:Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-se
 C:Genetics:
 A:Gene: aprI
 C:Superfamily: subtilisin homology
 C:Keywords: hydrolase
 F:1-40/Domain: signal sequence #status predicted <SIG>
 F:41-150/Domain: amino-terminal propeptide #status predicted <AMP>
 F:151-496/Product: alkaline serine protease I #status predicted <MAT>
 F:182-452/Domain: subtilisin homology <SBT>
 F:457-715/Domain: carboxyl-terminal propeptide #status predicted <CTP>
 F:239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 8.7%; Score 268.5; DB 2; Length 715;
 Best Local Similarity 25.8%; Pred. No. 3.5e-11;
 Matches 142; Conservative 51; Mismatches 201; Indels 157; Gaps 28;

QY 168 DFKQXNKEVOLRGIEIXAQQXXSNDVXYIT-----AKPEYVMND-----VARGIV 213
 Db 117 DKSLAKEQAQAFNEVVA---SGNVEYIEIDQMLKFPAPNDPRYDQWYEAAGI- 171
 QY 214 KADVAQSSYGLYGOGQIYVAVDG---LDTRGRDSSMHEAFRGKITALYALGRTNAND 269
 Db 172 NAPAAMDK--ATGGGVVAVLDITGRPHLDADAILPEYDITSTFYVANDGAGANDARD 229
 QY 270 ---TNG-----HGTNVAG---SVLNGXKTNKGAPOANLVFQS 301
 Db 230 PGDAVTRGEGCTDSSGQVPVRADQDSSWHGTHVAGTAAVTNNGEGVAGVADAKVVR 289
 QY 302 IMDSXGLIGLPSNLQTLFSAQASAGARIHTNSGAAV-----NCAYTTDSRNVDDY 354
 Db 290 VL---GKCGGLTSDIAGIILWASGSDRVPANAPAVYINNSLGGGACSAITQMAINO 346
 QY 355 RKNMTILFAAGNEXPNCTISAPGTAKNATVGAT--ENLRPSFGSYADNINHVAQESS 412
 Db 347 RNNGTVIYIAGNDNDSANTN-PGNCNGVYVAVSVDGSRATYSNGANI----- 397
 QY 413 RGPTRKDGRIKPDVNAPTGTXILSARSSLAPDSSFVANHDS-----KYAVMGTSNATP 464
 Db 398 -----DVAAPG---GAQSFADDPGILSTHNSGSGAPSDNSYHSQGTSMAP 442
 QY 465 IVAGNVQLR-----EHFVKN--RGITRKPISLRAAL--AGAADXGIG-----Y 505
 Db 443 HVAQVAALIKQAKSPAPDEYETILKNTTRSFAGSCGNCGTGVDAAVAANEALGVDVTP 502
 QY 506 PNCNOGWRVTLDSKLNVAAVNESSXLSOKATYXFTATAGKPLKLSLWSDAPASTTA 565
 Db 503 PTGN-----TLED--GVAKTGLSGAAGSND--FTTDDVRAK-----TNV 538
 QY 566 SVTL---VNDLDVITAPNGTYVGNDFXPPXXNMGR---NNVENVFINPQSGTYT 618
 Db 539 TFTWSGGTGADADLYVK-----LGSG---PTSSSYDCRPYEGGNALVCSFADPAQACTYH 588
 QY 619 IEVQATVPVPG 629
 Db 589 VMINGYKAYSG 599

RESULT 12
 F69730
 cell wall-associated protein precursor wprA - Bacillus subtilis

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:28:13 ; Search time 9.99218 Seconds

2652.411 million cell updates/sec

Title: US-09-920-954-1
Page: 3101

Sequence: 1 MRKKKVFSLSAAILSTV.....EVQAYNPVGPQXFLAIVN 6399

Scoring table: BLOSUM62

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum	DB	seq	length:	0
Maximum	DB	seq	length:	3

Maximum DB seq length: 200000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_40:*

Pried. No. Is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	518.5	16.7	1743	1	TAGC_DTCDD	Q2368 ditryostel
2	477.5	15.4	1305	1	TAGC_DTCDD	P54693 ditryostel
3	299.5	9.7	1398	1	TRGS_PYPYR	P27186 pyrococcus
4	286	9.2	806	1	SRBY_BACSU	P29141 bacillus s
5	274.5	8.9	420	1	SRBY_BACSU	P28842 bacillus s
6	273.5	8.8	580	1	PROA_XMACP	P23314 xanthomonas
7	262	8.4	694	1	WFA_BACSU	P54423 bacillus s
8	257.5	8.3	578	1	ELYA_BACSP	P20724 bacillus s
9	255	8.2	381	1	SQBL_BACNA	P35835 bacillus s
10	255	8.2	513	1	AQBL_THBAQ	P08594 thermus aq
11	253	8.2	382	1	SRBT_BACAM	P00782 bacillus a
12	251.5	8.1	381	1	SRBT_BACAM	P00783 bacillus a
13	250.5	8.1	380	1	ELYA_BACAO	P27693 bacillus a
14	250.5	8.1	380	1	ELYA_BACSU	P41362 bacillus a
15	249.5	8.0	379	1	SRBT_BACLI	P00780 bacillus s
16	249.5	8.0	381	1	SRBT_BACLI	P00489 bacillus s
17	247.5	8.0	381	1	SRBT_BACST	P29142 bacillus s
18	246.5	7.9	402	1	ALP_CEPAC	P22918 cephalospo
19	239.5	7.7	401	1	THEB_BACSP	Q45670 bacillus s
20	239.5	7.7	534	1	PROA_VIBAL	P16588 vibrio alg
21	239	7.7	269	1	SUBS_BACSP	P29640 bacillus l
22	238	7.7	269	1	PRTM_BACSP	Q99405 bacillus s
23	234.5	7.6	1181	1	SCAA_STRRP	P29809 streptococ
24	232	7.5	269	1	SRBB_BACLE	P15926 streptococ
25	231.5	7.5	1167	1	SCAL_STRRP	Q93778 magnaporthe
26	223.5	7.2	404	1	SMPI_MGPO	O01420 trichoderma
27	223.5	7.2	409	1	ALP_TRHA	P29140 bacillus c
28	223	7.2	321	1	ISF_BACSS	P07518 bacillus p
29	218.5	7.0	275	1	SRBT_BACFU	P33285 aspergillus
30	218.5	7.0	533	1	PEPC_ASGNG	P41363 bacillus h
31	216.8	7.0	361	1	ELYA_BACMD	P16396 bacillus s
32	215.5	6.9	645	1	SRBE_BACSU	P29139 penicillium
33	214.5	6.9	326	1	ISF_PAEBO	

ALIGNMENTS

RESULT 1	TAGC_DICDI	STANDARD:	PRT: 1743 AA.
AC	Q23868;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Prestalk-specific protein tagc precursor (EC 3.4.21.-)		
GN	TAGC.		
OS	Dicystostellum discoidesum (Slime mold).		
OC	Eukaryota; Mycetozoa; Dicystostellidae; Dicystostellum.		
OX	NCBI_TaxID=44689;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SFRAN-AX4;		
RX	MEDLINE=97140317; PubMed=8986798;		
RA	Shaulsky G, Escalante R, Iocomis W.F.;		
RT	Developmental signal transduction pathways uncovered by genetic suppressors.		
RL	Proc. Natl. Acad. Sci. U.S.A. 93:15260-15265(1996).		
CC	-1- FUNCTION: INTERCELLULAR COMMUNICATION VIA TAGC MAY MEDIATE INTEGRATION OF CELLULAR DIFFERENTIATION WITH MORPHOGENESIS (BY SIMILARITY).		
CC	-1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO PEPTIDASE FAMILY S8.		
CC	-1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.		
CC	-1- SIMILARITY: STRONG, TO TAGB.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on ways that can be used by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).		
CC	-----		
CC	EMBL: U60086; AAB0331.1; -.		
DR	DicYdb; DD02060; tagc.		
DR	InterPro: IPR003593; AAA_Arpase.		
DR	InterPro: IPR003439; ABC_transport.		
DR	InterPro: IPR001140; ABCtransprtm.		
DR	InterPro: IPR000209; Peptidase_S8.		
DR	Pfam: PF00005; ABC_tran; 1.		
DR	Pfam: PF00664; ABC_membrane; 1.		
DR	PRINTS: PS00723; SUBTILISIN.		
DR	SMART: SM00382; AAA; 1.		
DR	PROSITE: PS00136; SUBTILASE_ASP; FALSE_NEG.		
DR	PROSITE: PS00137; SUBTILASE_HIS; 1.		
DR	PROSITE: PS00138; SUBTILASE_SRP; FALSE_NEG.		
DR	PROSITE: PS00211; ABC_TRANSPORTER; 1.		
KM	Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane; Signal.		
KT	1	27	POTENTIAL.
KT	28	1743	PRESTALK-SPECIFIC PROTEIN TAGC.
KT	316	642	PROTEASE.
KT			

```

FT DOMAIN ? 1743 ABC TRANSPORTER.
FT TRANSMEM 962 982 POTENTIAL.
FT TRANSMEM 1027 1047 POTENTIAL.
FT TRANSMEM 1072 1092 POTENTIAL.
FT TRANSMEM 1157 1177 POTENTIAL.
FT TRANSMEM 1260 1280 POTENTIAL.
FT TRANSMEM 1288 1308 POTENTIAL.
FT ACT_SITE 325 325 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 372 372 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 637 637 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT NP_BIND 1485 1492 ATP (POTENTIAL).
FT DOMAIN 42 46 POLY-ASN.
FT DOMAIN 94 103 POLY-ASN.
FT DOMAIN 643 646 POLY-ALA.
FT DOMAIN 733 741 POLY-ASN.
FT DOMAIN 786 792 POLY-SER.
FT DOMAIN 1337 1340 POLY-GLU.
FT DOMAIN 1346 1352 POLY-GLY.
FT DOMAIN 1353 1357 POLY-ASN.
FT DOMAIN 1358 1364 POLY-ASP.
FT DOMAIN 1381 1386 POLY-ASN.
FT DOMAIN 1707 1729 POLY-ASN.
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 547 547 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 614 614 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 689 689 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 735 735 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 741 741 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 776 776 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 832 832 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1251 1251 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1385 1385 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1386 1386 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1454 1454 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1704 1704 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1743 AA; 194145 MW; 12DB363E2F729839 CRC64;

```

Query Match 16.7%; Score 518.5; DB 1; Length 1743;
 Best Local Similarity 24.2%; Pred. No. 4.1e-30;
 Matches 194; Conservative 114; Mismatches 213; Indels 281; Gaps 33;

```

QY 87 HIXQFNGPILETRKQX-----EXTGAKIDYIPDAVIV---EYEGDVXSK 130
DB 115 YIVQFKRIDNDETRQLEFLIGTDIVDEQYOSHIVYIPHDSFLVLMQEGSVLLSS 174
QY 131 XXXIEHVSVEPY-----LPXYXIDPOLFTKGASXLVK--AXALDT--KQXN 173
DB 175 KEWMSWIGEFEPSSNKHILNNEKSGIGLPYII---LSDSTNSLIQREWENTLSLKSYN 230
QY 174 KEVOLRGI-----EXIAQXXXSNDVXYITTKKPEYK 203
DB 221 SKVKLITLINOXKLSIYVNCNDESPSCSLINSEKLYIOWISEQESN---FIENSEKFQ 287
QY 204 VMNDVARGIV-----KADVAOSSGYLGOGQIVAAVADGLDGR--NDS----- 245
DB 288 TANRISPRVVGCTKDTLVNNDVRDIP-----LRGKGQILSTADTGLDSHCFQFSKXPI 342
QY 246 ---SMHEAFKRTALYALGRTNNAANDTNGCHTHVAGSVLG-----NGXTKNMADQA 295
DB 343 PLNSVNLNHR-KVYVYITSTSDSDSKVDGCHGTHICGSAAGTPEDSSVNISFSGLATDA 401
QY 296 NLVPSIMDSXGGLGL--PENIQLTFQOAXSAGARIFTNMGW-----AVNGAYITDSRN 349
DB 402 KIAF--FDLASGSSSLPPSDLKQLYOPLYDAGARVACDSWGSVEGYVSSDPAAS 458
QY 350 VDDYVRKN-DMTILFAAGNEXPNGTIS--APGTAKNAITGATENLR-----PSFGS 399
DB 459 IDDFLFTHPDITILRAAN---NBOYLSLLQSTAKNAVITYGAGHTIHENLITDGPNTIN 515
QY 400 YADNT----- 404

```

```

DB 516 YQSSVDINQELICDPDSRYCNYTTAQCCLSNATTTGLASCPILLKRSVIDAANTOPLY 575
QY 405 --NHVAFSSRGPTKDRIKPDVNAFGTXILSARSSLA-----PDSSFMANHDSKYAYM 456
DB 576 NENNICSSGCPHDDRMKRALVAPGEYITSAANSANANTDQCGDSL--PNTALLA-I 633
QY 457 GGTSMAPIYAGNVAQLREH-----FVKRGITPRPSLKALALIGA----- 498
DB 634 SGTSMASFFAAMATTLROYLVADGYPTGSIVESNKLQPGSLKALMINNAQLLNGTFQ 693
QY 499 --AXXGIGYNGN-----QGMGRVTLDSLVNAYVNESS----- 530
DB 694 LITSSSITPBNQVFNENFACASLVQMGALIMSMWLVHVNNNNNNNKTSDDGTRPDGI 753
QY 531 -----XLTSSQKATYFT-----ATAGKPLK--ISLWSDAPASTT 564
DB 754 GGLDLRLVKPMQKKEESLSTQNTSYCFYTPSSSSNSGNNIPRVATLVWTFPPSYAG 813
QY 565 ASVTLVNDLIVT-----TAPNGTYXGNDPFXAXXNMGRNVENVFTINXP 612
DB 814 AKFMLVNNLDTLMTYRDNGSTIFYSNQGSFFLG-----LAPTDPLNNVEGIVHNPT 867
QY 613 QSGTYTIEVQAYNPVGPQXFS 634
DB 868 EPMTRYEMWACTNVPMPQPNFS 889

```

RESULT 2

TAGB_DICDI STANDARD; PRT; 1905 AA.

ID TAGB_DICDI

AC P34683;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Prestalk-specific protein tagB precursor (EC 3.4.21.-).

GN TAGB.

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

NCBI_TaxID=44689;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AX4;

RA MEDLINE=95262903; PubMed=7744252;

RA Shaubsky G., Kuspa A., Loomis W.F.;

RT "A multidrug resistance transporter/serine protease gene is required

for prestalk specialization in Dictyostelium.";

RL Genes Dev. 9:1111-1122(1995).

CC -1- FUNCTION: INTERCELLULAR COMMUNICATION VIA TAGB MAY MEDIATE

INTEGRATION OF CELLULAR DIFFERENTIATION WITH MORPHOGENESIS.

CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO PEPTIDASE FAMILY

S8.

CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING

TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.

CC -1- SIMILARITY: STRONG, TO TAGC.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC EMBL: U20432; AAA62212.1; -.

DR MEROPS: S08.0PW; -.

DR Dictydb: DD02059; tagB.

DR InterPro: IPR003593; AAA_ATPase.

DR InterPro: IPR003439; ABC_transporter.

DR InterPro: IPR001140; ABCtransprtTM.

DR InterPro: IPR000209; peptidase_S8.

DR Pfam: PF00005; ABC_tran.1.

DR Pfam: PF00664; ABC_membrane.1.

DR PRINTS: PR00723; SUBTILISIN.

```

DR ProDom: PD000006; ABC-transport: 1.
DR SMART: SM00382; AAA: 1.
DR PROSITE: PS00136; SUBTILASE-ASP; FALSE-NEG.
DR PROSITE: PS00137; SUBTILASE-HIS; 1.
DR PROSITE: PS00138; SUBTILASE-SER; 1.
DR PROSITE: PS00211; ABC-TRANSPORTER; 1.
DR Hydrolyase: Serine protease; ATP-binding; Transport; Transmembrane;
KW Signal.
FT SIGNAL. 1 31
FT CHAIN 32 1905
FT DOMAIN 378 700
FT DOMAIN 1011 1031
FT TRANSMEM 1076 1096
FT TRANSMEM 1121 1141
FT TRANSMEM 1210 1230
FT TRANSMEM 1309 1329
FT TRANSMEM 1332 1352
FT AC1_SITE 387 387
FT AC1_SITE 432 432
FT AC1_SITE 432 432
FT AC1_SITE 695 695
FT NP_BIND 1553 1560
FT DOMAIN 63 67
FT DOMAIN 95 104
FT DOMAIN 107 134
FT DOMAIN 311 321
FT DOMAIN 833 837
FT DOMAIN 838 844
FT DOMAIN 871 876
FT DOMAIN 1012 1015
FT DOMAIN 1386 1389
FT DOMAIN 1445 1450
FT DOMAIN 1765 1779
FT DOMAIN 1782 1785
FT DOMAIN 1807 1812
FT DOMAIN 1813 1860
FT DOMAIN 1872 1878
FT CARBOHYD 594 594
FT CARBOHYD 621 621
FT CARBOHYD 672 672
FT CARBOHYD 747 747
FT CARBOHYD 823 823
FT CARBOHYD 1172 1172
FT CARBOHYD 1522 1522
FT CARBOHYD 1658 1658
SQ SEQUENCE 1905 AA; 212518 MW; B8E23FAB89AE13C CRC64;

Query Match 15.4%; Score 477.5; DB: 1; Length 1905;
Best Local Similarity 24.5%; Pred. No. 4; 8e-37;
Matches 193; Conservative 98; Mismatches 237; Indels 259; Gaps 31;

OY 87 HIXFNGPILEETKXLE-----XTGAKILDYIPDYAVIYER-----BGDYXS 129
DB 172 YIVQPKRIDETREQLKEFLIGTITLLKQPKRSHIVHIFHDSFLVFMTRKQSVLLS 231
OY 130 XXXXIEHVESVEPY-----LPXXYI-----DPQLFTKGASXL-----V 162
DB 232 SKEMISWIGHEPSNKHILNHYEKSIGYPVITILSGSTNSLIQRMWNTLSILTSYNSKV 291
OY 163 KAXALDTRQ-----XNKEVOLRGIEIXIAQXXXS--NDVYXITAKPEYVM 205
DB 292 KLTLINRKLIKSLIYCNDESPSSSSSCSLIGSEKIVYKWISEGESNVIERSEKLTQA 351
OY 206 NDVARGIV-----KADVAQSSYGLYGOGQIYAAVADTGIDPCR--NDS----- 245
DB 352 NRLSPFVIFCTKDKLVNNDRIDIP-----LRKGQGLIISINDTGLDSDSHCFDSKYPPIF 406
OY 246 -SMHEAFRGKITALVLTGRTNNAANDTNGHGVAGSVLGNCXN-----KGMAPQANL 297
DB 407 NOVENENRKRYVT--YITHDNEVDVYNGHGHVCGSAAGPEPDSNAISSGSLTADKTI 463
OY 298 VFOGIMSDXGGLGGLPSNLQTLFSAQXASGARIFHTNSWGA-----AVNGAYTTDSNRVDY 353

```

```

DB 464 AFYD-LSSGSESEPTPEPDSQMYKPELYDAGARVHGDSWGSVSLGGYGGSDDAGGIDAF 522
OY 354 VRK-NDMTILFAAGNEXPRGCTISAPGAKNAIYVGAETELRLRPSFGS-----YADNI- 404
DB 523 LYEPERSTLRAAGNN-ELFASFLAQATAKNAIYVGAEOFAHNVYSALLEYDFPSDMAN 581
OY 405 -----NHVAOF 410
DB 562 FQRCLEFDKRYCYTTAKCCSEVSNVKGLOLCCPASTIKONASDFTTORPFINENMKSF 641
OY 411 SSRGPTKDCGIRKVDWAPGCTYXLSARRS-----LAPDSFMANHDSKYAYGCTSNATP 464
DB 642 SSKGPTHGKLDKPDVIAPEYITTSARSGENSTDOCGDSL--PNANGIMLSIGTSMATP 699
OY 465 IVAGNVAQLREHF-----YKNRGITPPKPSLLKALII----- 495
DB 700 LATPATILRQYLVLDGFPPTGESVEENKILPTGSLKALMINNAOLLNGYFWSASSTNP 759
OY 496 AGAADKGLGYPNCGMGHVTLDKSLNVAAYNESS-----XLSOKAT----- 539
DB 760 SNAIFEDINGANLIQGWGALRMN--NMLYKSSNPPTPSRWIGIGLGCKNOKATEWKED 816
OY 540 -----YXFT-----ATACKP-LKISLWSDAPASTASTATVYNDLTL-- 575
DB 817 SLSSGLNKSCYCFYKRPSSSSSGSGGCTPRIVATLVMTDPPSYSAKFNLYNNDLILL 876
OY 576 -----YITAPN--GTYHYGNDFFXXPPXXMMDGNVENYFINXQSGTYTLEVOAYNP 627
DB 877 NSDDDSITITGNSGSLQPAKVAP-----DTLVNVEGIIITPKANNYKFTTAGTNVP 931
OY 628 VGPQXFS 634
DB 932 IGPQKFS 938

RESULT 3
PLUS_PYREFU STANDARD; PRT: 1398 AA.
ID PLUS_PYREFU
AC P7186;
DT 16-OCT-2001 (Rel. 40, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pyrolysin precursor (EC 3.4.21.-).
GN PLS OR PF0287.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 150-184, AND CHARACTERIZATION.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RX MEDLINE=96355370; PubMed=8702780;
RA Voorthorst W.G.B., Eggen R.I.L., Geerling A.C.M., Platteeuw C.,
RA Slezan R.J., de Vos W.M.;
RT "Isolation and characterization of the hyperthermostable serine
RT protease, pyrolysin, and its gene from the hyperthermophilic archaeon
RT Pyrococcus furiosus.";
RL J. Biol. Chem. 271:20426-20431(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RN Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
RX [3]
RP CHARACTERIZATION, AND 3D-STRUCTURE MODELING.
RX MEDLINE=21079021; PubMed=11210516;
RA de Vos W.M., Voorthorst W.G.B., Dijkgraaf M., Kluskens L.D.,
RA van der Oost J., Slezan R.J.;
RT "Purification, characterization, and molecular modeling of pyrolysin
RT and other extracellular thermostable serine proteases from
RT hyperthermophilic microorganisms.";

```

```

RL Meth. Enzymol. 330:383-393(2001).
CC -1- FUNCTION: Has endopeptidase activity toward caseins, casein
CC fragments including alpha-S1-casein and synthetic peptides.
CC -1- SUBCELLULAR LOCATION: Cell-envelope associated.
CC -1- PFM: LWM pyrolysin seems to be produced by autoprotoeolytic
CC activation of HWM pyrolysin.
CC -1- PFM: Glycosylated.
CC -1- MISCELLANEOUS: Thermostable; high activity at 95 degrees Celsius.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DR EMBL, U55835; AB09761.1; -.
DR EMBL, AE010153; ALB0411.1; -.
DR HSSP; Q45670; 1DB1.
DR MEROPS; S08.100; -.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 4.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KM Hydrolase; Serine protease; Zymogen; Signal:
KM Complete proteome.
FT SIGNAL 1
FT PROPEP 1
FT CHAIN 1
FT ACT_SITE 150 1398
FT ACT_SITE 179 179
FT ACT_SITE 365 365
FT ACT_SITE 441 441
FT CARBOHYD 152 152
FT CARBOHYD 222 222
FT CARBOHYD 228 228
FT CARBOHYD 240 240
FT CARBOHYD 257 257
FT CARBOHYD 262 262
FT CARBOHYD 298 298
FT CARBOHYD 327 327
FT CARBOHYD 406 406
FT CARBOHYD 651 651
FT CARBOHYD 663 663
FT CARBOHYD 739 739
FT CARBOHYD 792 792
FT CARBOHYD 893 893
FT CARBOHYD 908 908
FT CARBOHYD 917 917
FT CARBOHYD 929 929
FT CARBOHYD 1048 1048
FT CARBOHYD 1056 1056
FT CARBOHYD 1084 1084
FT CARBOHYD 1117 1117
FT CARBOHYD 1133 1133
FT CARBOHYD 1140 1140
FT CARBOHYD 1148 1148
FT CARBOHYD 1208 1208
FT CARBOHYD 1233 1233
FT CARBOHYD 1237 1237
FT CARBOHYD 1332 1332
FT CONFLICT 607 609
FT CONFLICT 881 881
SO SEQUENCE 1398 AA; 154474 MW; 355D873A27D56552 CRC64;
Query Match 9.7%; Score 299.5; DB 1; Length 1398;
Best Local Similarity 22.9%; Pred. No. 3.9e-14;
Matches 173; Conservative 76; Mismatches 248; Indels 257; Gaps 31;
QY 96 LEETKXOEXETAKILID-----YIPDIAYIVEY-----GDVXSKXXXTIE 135

```

```

Db 77 LEEAKTELEKLAELLIDENRVLNMLLKIKPERKVELNITSSLEKAMLRREVLSPPYE 136
QY 136 -HVESVEPLPYXIDDPOLFT-----KGASXVKKAKALPTKONKEV 176
Db 137 KDKTKEP-----SLEPKRYNSTWYNALQFOEGCYDSSGVVAVLDTGVDPNHPFLSI 191
QY 177 QLRGIEIXIAQXXXNDVXYITAKPEY-KVM-----DVARGI----- 212
Db 192 PFDGRKKILEMKDFDEGEVDTSFSPSKVYNGTLINTTFQVAGSLTNESTGLMEYVK 251
QY 213 -----VKADVAAOSSGLY-----GGOQIYAVADT 236
Db 252 TVYVSNVTIGNITTSANGIYHFGILPERKPDLPDGDQEDFYVLLVNSTGNYDIAYVDT 311
QY 237 GLDTGRNDS-----SMHEAFRGKITALYALGRTNAN-----DTNGHGTNVAG 279
Db 312 DLDYDFTDEVPLAQYVNTYDVAVFSTYTGPLNVLAEIDPNCETAVFGDHGHTNVAG 371
QY 280 SVLGNXGXTN-----KGMAQANLVFQSIMDSXG 307
Db 372 TVAGYDSNNDDAMDWLSMTSGEMEVEFSLYGMWDTNTYDTGVGVAPGAQIMALIVLRS-D 430
QY 308 GLGLGPSNLQTLPSQAXSAGARIHTNSMGAVNAVY--TDSRN--VDDYVRKNDMTILF 363
Db 431 GRGSMWDILEGM-TYAATGADVISMISLGG--NAPYLDGTDESYAVDELTEKYGVFEVI 487
QY 364 AAGNEKPNNGTISAPGTANATTCATENTLRSPFSYAD-----NTN 405
Db 488 AAGNEGPNGINIVSGPVAKKALTGVGA--VPINVGYYVSOALGYPDYIGFYFPAYITN- 545
QY 406 HVAQFSSRGPTKGRIPKPVMAPGTXILSARSSLAPODSFMANHDSKYAVMGTSNAPPI 465
Db 546 RIAPFSSRRPRIDGEIKRPNVAVAPGYISLSLMMWGGAF-----MSGTSMATPH 595
QY 466 VAGNVAQLEHNVKRNRTPKPSLKAALIAQA-----DXGLGYPNGNCGWGTLDK 519
Db 596 VSGVVALLLTSG-AKEAGCIYNNPDIIRKVLSEGATWLEGGDPYTGQRYTELDOHGLVNYK 654
QY 520 SLNVAVYVNSSKXLSQKATYFTATAGKPLKILSYLSWAPASTTA---SVTLNDDLV 576
Db 655 SWEI-----LKAINGTTLPLVDHMDKSIDAEYIGVYINGLAR 696
QY 577 ITAPN-----GTYVYVG---DEXXPXXNMW-----DG-----RNVE----- 605
Db 697 NSIPDIVEHNIKVYGDTEKRETEIYATBPMIKPFVSGVILNNTEFVLRYKDYEGLEP 756
QY 606 -----NVFINXQSGITYTTEVOAYNVPGQXFS 634
Db 757 GLYVGRILIDPT--TPVIEDEILNTIVIPKFT 788
RESULT 4
SUBV_BACSU STANDARD; PRT; 806 AA.
AC P29141;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Minor extracellular protease Vpr precursor (EC 3.4.21.-).
GN VPR OR IPA-45R.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 161-195.
RX MEDLINE=92041574; PubMed=1938892;
RA Sloma A., Ruto G.A. Jr., Theriault K.A., Dwyer M., Wilson S.W.,
RA Pero J.;
RT "Cloning and characterization of the gene for an additional
RT extracellular serine protease of Bacillus subtilis.";
RL J. Bacteriol. 173:6889-6895(1991).
[2]

```

RP SEQUENCE FROM N.A.
RC STRAIN-168:
RX MEDLINE-95020537; PubMed-7934828;
RA Glaser P., Kunst F., Arnaut M., Coudart M.P., Gonzales W.,
RA Hillo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
RA Prescan E., Santana M., Schneider E., Schweizer J., Vertes A.,
RA Rapoport G., Danchin A.;
RT "Bacillus subtilis genome project: cloning and sequencing of the 97
RT kb region from 325 degrees to 333 degrees.";
RL MOL. Microbiol. 10:371-384(1993).
RP SEQUENCE FROM N.A.
RC STRAIN-168:
RX MEDLINE-98044033; PubMed-9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Allon G.,
RA Azevedo V., Bartori M.G., Bessieres P., Bolotin A., Borczyk S.,
RA Borris R., Bourlier L., Brans A., Brann M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano N.J., Carter N.M.,
RA Choi S.R., Codan J.W., Conneton I.F., Cummings D., Daniel R.A.,
RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ertlan K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hillbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Oglwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetle D., Portolick S., Prescott A.M.,
RA Prescan E., Pulic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche M., Sadle Y.,
RA Sato T., Scanlan E., Schleich S., Schoeter R., Scoffone C.,
RA Sakuchi J., Sekowska A., Seror S.J., Shin B.S., Soldo B.,
RA Sorochin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpsita P., Tognoli A.,
RA Tosato V., Uchiyama S., Vandebol M., Vanlier F., Vassarotti A.,
RA Viari A., Mamult K., Wedler E., Wedler H., Weizenegger T.,
RA Winters P., Wipal A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*.
RL Nature 390:249-256(1997).
CC -1- FUNCTION: NOT REQUIRED FOR GROWTH OR SPOULATION.
CC -1- SUBCELLULAR LOCATION: secreted.
CC -1- PTM: PROBABLY UNDERGOES C-TERMINAL PROCESSING OR PROTEOLYSIS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M76590; AAA22881.1; -;
DR EMBL: X73124; CA51601.1; -;
DR EMBL: Z99123; CAB15835.1; -;
DR PIR: A41341; A41341.
DR HSSP: P00782; 2SPT.
DR MEROPS: S08.00A; -;
DR Subtilist; BG10591; vpr.
DR InterPro: IPR003137; PA.
DR Pfam: PF00082; Peptidase_S8; 2.
DR Pfam: PF02225; PA; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.

KW Hydrolase; Serine protease; Zymogen; Signal; Complete proteome.
FT SIGNAL 1 28
FT PROPEP 29 160
FT CHAIN 161 806 MINOR EXTRACELLULAR PROTEASE VPR.
FT ACT_SITE 169 189 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 233 233 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 534 534 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 806 AA; 85608 MW; F984E3BF0869DDDC CMC64;
Query Match 9.28; Score 286; DB 1; Length 806;
Best Local Similarity 22.58; Pred. No. 18e-13;
Matches 132; Conservative 60; Mismatches 188; Indels 206; Gaps 19;
QY 155 TRKASLVYAAAXLIDTKQAKVEQ--LGIT-----EXIQAXXSNDDYVITAKREYVNM 206
DB 87 TRKAKNAIK--AVKNGKAKREYEVSGFSMKLPAHEIKLLAVDVAVPPNTYTDN 144
QY 207 DVARGIVKADVAQS-----SYLGQGVAVVADTGLDTCR----- 242
DB 145 MKDKDVTISEDVAVSQMDSDAPYIGANDAMDLDYTGKGIKVAIIDTGVYHHPDLKKNFG 204
QY 243 -----NDSSMHEAFRGKITALVALGRTNNANDNGHGVAVSVLGNQXTKGMAP 293
DB 205 QYKGYDFVNDYDPEKTEPTG-----DPRGEATDHGTHVAGTVAAANG--TIKGVAP 252
QY 294 QANLVFQSIIMDSXGLGLSPENLQTLFSAQXAGARIHTNSGAAVNGAYTTDSRNVDDY 353
DB 253 DATLLAVRYLIGFG--SGTTEENVLAGVERAVQDGADVMNLISGLNSLNPDMVATSTAL-DW 309
QY 354 VRKNDMTIFFAAGNEXPMGGSITASGRTAKNAIVGATE-----NLNRPFGSY----- 400
DB 310 AMEGGVAVTSGNNGSPRGWTVGSGPSTRSAISVGTQLPLNEVAVTVGTSISAKVGVYN 369
QY 401 -----ADNI----- 404
DB 370 KEDDYVALNKKEVELVEAGIEAADFEGCKDLTKVAVVVRGSIADVADNAKKAAGIAG 429
QY 405 -----NHVAQF 410
DB 430 VYNNLSGEIEANVPGMSVPTIKLSLEDGKLVSAALKAGETTKTKLVSKALGEQVADF 489
QY 411 SSRGPKRD-GRIKPDVMAPGTYILSARSSLAPOSSFWAMHDSKYAVMGCTSMATPIVAGN 469
DB 490 SSRGPMVDMTMIKIPDISAPGVNIIVSTPTHPD-----HPYGYSGKGTSMASPHIAGA 543
QY 470 VAOLREFFVKNRGIPTKPSL--LKAALIAAGAA-----DXGLGYPNGQMGCVTLDKSLNV 523
DB 544 VAVIKQ-----AKPKMSVEQIKAAIMNTAVTLKSDSEVYPRHNAQAGSARI--MNA 593
QY 524 AYVNESSXLSTSQKATYXFTATPAGKPKISLVMSDAPASTASVTL 569
DB 594 --IKADSLVSPGSYGVGFLEKNENETKNTFTIEQSSIRKSYTL 637
RESULT 5
SUBT_BACS9
ID SUBT_BACS9 STANDARD: PRT; 420 AA.
AC P28842;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Subtilisin precursor (EC 3.4.21.62).
GN SUBI.
OS *Bacillus* sp. (strain TA39).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=29336;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92256481; PubMed-1581352;
RX Martin E., Davail S., Feller G.;
RT "Nucleotide and derived amino acid sequence of the subtilisin from
RT the antarctic psychrotroph *Bacillus* TA39.";
RL Biochim. Biophys. Acta 1131:111-113(1992).

```

CC -1- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE,
CC IF CATALYZES THE HYDROLYSIS OF PROTEINS AND PEPTIDE AMIDES.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in pI. Hydrolyzes peptide amides.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: STILL ACTIVE AT TEMPERATURES CLOSE TO 0 DEGREES
CC CELSIUS, IT HAS A MARKED HEAT STABILITY.
CC -1- MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF
CC SPORE germination AND MANY MUTATIONS WHICH BLOCK SPORE germination AT EARLY
CC STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN
CC IS NOT NECESSARY FOR NORMAL SPORE germination.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X62369; CAA44227.1; -
DR PIR: S23407; S23407.
DR HSSP: Q99405; 1MPT.
DR MEROPS: S08.0PA; -.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASASP; 1.
DR PROSITE: PS00137; SUBTILAS_HIS; 1.
DR PROSITE: PS00138; SUBTILAS_SER; 1.
DR HydroLase: Sporulation; Serine protease; zymogen; signal.
FT SIGNAL 1 26 POTENTIAL.
FT PROPEP 27 111 POTENTIAL.
FT CHAIN 112 420 SUBTILISIN.
FT ACT_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 182 182 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 360 360 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 420 AA; 44086 MW; A4F121BD32826EC CRC64;

Query Match 8.98; Score 274.5; DB 1; Length 420;
Best Local Similarity 28.38; Pred. No. 5.4e-13;
Matches 107; Conservative 56; Mismatches 142; Indels 73; Gaps 21;

QY 167 LDTRKXN--KEVQLRGIXIAQXXSNDVXYITAKPE--YKVM---NDVARGIV---KA 215
DB 73 MNEQFNLKKNKNTLVKPEL---EIAATDKPEALYNMAASOSTPWCIRAIYNNNS 128
QY 216 DVAOSSYGLYQGOIVAVADTGLDTRGNDSSMHEAFKRGKITAL--YALGRT--NNANDT 270
DB 129 SITQTS---GGGGINIAVLDTGVNTN-----HPDLRNNEQCKDFVTGTYTNNSCDR 179
QY 271 NGHGHVAVSVLNGXNTRK---GNAPQANLVFGISIM--DSKGGGLGLPSNLOTLEFSQXS 325
DB 180 QGHGHVAVSGLADGCTNGVYGAVPADLWAVYKVLGDGSGSYADIDIAALRHAGDQATA 239
QY 326 AGARHTN--SWGAVNGAYTTDSRVNDYVRKNDMTIFLAAGNEXPCNGCTISAPGTAKNA 384
DB 240 LNTGVIMTSLGSGESSLITNNAV---YSYNNKGLIITAAAGNSPFGSGISYCPALVNA 296
QY 385 ITVGATENTLRPSFGSYADINHHVAOFSSRGP--KDG---RIKPDVMAFGTYILSARSS 438
DB 297 VAVVALEN--KYENGTY-----RVADFSRGRYSMTGDGAIVKQGDVEISAPGAIAYST--- 347
QY 439 LAPDSSFPANNDKAYAVNGTSMATPIYAGVNAOLREHFVANKRGITPPPSL-----L 490
DB 348 -----W--FDGGVATISGTSMASPHAAGLAKIMAOYPSASNDVGELOYRAYENDI 398
QY 491 KAALTAGAAD--XGIGY 505
DB 399 LSGYIAGYGDFASGFGF 416

```

```

RESULT 6
PROA_XANCP
ID PROA_XANCP STANDARD: PRT: 580 AA.
AC P23314;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Extracellular protease precursor (EC 3.4.21.-).
GN XCC0851.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90251253; PubMed=2187155;
RA Liu Y.-N., Tang J.-L., Clarke B.R., Dow J.M., Danlels M.J.;
RT "A multipurpose broad host range cloning vector and its use to
RT characterise an extracellular protease gene of Xanthomonas campestris
RT pathovar campestris."
RL Mol. Gen. Genet. 220:433-440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reznach F.C., Farah C.S., Furian L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Clapina L.P.,
RA Claretelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X51635; CAA35962.1; -
DR EMBL: AE012184; AAM40166.1; -
DR PIR: S11890; S11890.
DR HSSP: P00782; 2SPT.
DR MEROPS: S08.0PA; -.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASASP; 1.
DR PROSITE: PS00137; SUBTILAS_HIS; 1.
DR PROSITE: PS00138; SUBTILAS_SER; 1.
DR HydroLase: Serine protease; zymogen; signal.
FT SIGNAL 1 32 POTENTIAL.
FT PROPEP 33 2136 POTENTIAL.
FT CHAIN 137 580 EXTRACELLULAR PROTEASE.
FT ACT_SITE 177 177 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 237 237 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 409 409 CHARGE RELAY SYSTEM (BY SIMILARITY).

```



```

FT DISULFID 225 273 BY SIMILARITY.
FT DISULFID 315 352 BY SIMILARITY.
FT DISULFID 450 454 POTENTIAL.
SQ SEQUENCE 580 AA: 57228 MW: 8C9A2CA4E7F47CB CRC64;

Query Match 8.8%; Score 273.5; DB 1; Length 580;
Best Local Similarity 22.7%; Pred. No. 9.8e-13;
Matches 159; Conservative 80; Mismatches 250; Indels 211; Gaps 30;

OY 1 MRKKVFLSYLSAALITVVALXNPSACXAXXFLDLKRGQIOTTDXXGFSKQXQXGAFAF 60
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 6 LKRTGSLTIGASALTSLLLMPAFAG-----EYIDGLATQHQKLTIVYKNGSIAL 60
OY 61 LVESENVKLGKGLKLETPPANNNLHXQNPILFETKQXLEHGACILIDYIDYAYI 120
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 61 ASPS-----ALTTSLRTAARAVPA-----KAGKALGINSVRLALGELY-- 100
OY 121 VEEYGDVXKXXXXIEHVESEVEYELFYXIDQLETKASXLYVAKAALDTKQXNKEVQLRG 180
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 101 -----RADRALDRAELETLMQLADPNVQSEV----- 129
OY 181 IETIAQXXXNDVXYITANP-EYKYMDVARGIVKA--DVAQSSYGLYGGQIVAVADTG 237
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 130 -----DQILHATLTPNDTRLEQOMAFGTNAGLINIRPAMDKATGSGTVAVIDTG 179
OY 238 L-----DTGRNDSMHEAFRGKITALYALGRNNAND-----TNGHGT 275
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 180 ITSHADLNANILAGYDFISDATTARDGNGRSDNADGDMYAAINECGAGIPAASSSWHGT 239
OY 276 HVAGSY--LGNQXTN-KGMAPQANLYFGSINDXSG-----LGLIPENL 316
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 240 HVATVVAAYNNNTGVAGTAAGAKVPRVYLGKGGSLSDIADAIYASGAGVSGITPANA 299
OY 317 Q--TLFQAASAGARHTMSQCAVNGAYTDSRNVDDYRRKDMTILFAAGNEPANGGT 374
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 300 NPAVYIMSLGGGSGSTTMQW-ALNGAVNGGT-----TVVAAGADSNVSG 346
OY 375 ISAGCTKKAATYTGATEN--LPSFGSADNINHVAFSSRGPTKQGRKIPVMAPGTYI 432
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 347 -SLPANCANIAVAATTSAGAAKASVNGGTGT-----DVSAPEGSSI 386
OY 433 LSNRSS--LAPDSSFWANHDSRYAVMGSTMATPIVAGNVAQLREHFVKNRGITPK--PS 488
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 387 LSTINSGTTPPGS-----ASVASYNGTSMASPHVAGVALVQS--VAPTLATPAAVET 437
OY 489 LLK--AALINGAADXGLGY-----PNCNGCMGRVTLDKSLNVAAYNESSX 531
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 438 LKNTATATLPGACGCGAGIYNADAVTAALINGSGGGGNTLNGTPTVGLG---- 493
OY 532 LSTQKATYFTATAGKPLKISLWSDAPASTASVTL--VMDLQVY--TAGNGTY 585
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 494 AATGAELNYITIVPAG-----SGTLIVTSSGSDADLIYRAGSAFTDSAY 539
OY 586 VGNDFXXXKXNMDGNNTNENFIKXPGSGTYTTEGVAYN 625
DB 540 TCRPIRS-----GNAETCTITAP-SGIYVRLKAYS 569

RESULT 7
WPA_BACSU STANDARD: PRT: 894 AA.
AC PS4423; 006726;
DR 01-OCT-1996 (Rel. 34, Created)
DR 15-DEC-1998 (Rel. 37, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell wall-associated protease precursor (PC 3.4.21.-) [Contains: cell
DE wall-associated polypeptides CMBP23 and CMBP52].
GN WPA.
OS Bacteria subtlis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54 AND 414-428.

```

```

RC STRAIN-168;
RX MEDLINE-97158234; PubMed-9004506;
RA Margot P., Karamata D.;
RT "The wprA gene of Bacillus subtilis 168, expressed during exponential
RT growth, encodes a cell-wall-associated protease.";
RL Microbiology 142:3437-3444(1996).
RP SEQUENCE FROM N.A.
RX MEDLINE-98015415; PubMed-9353931;
RA Medina N., Vanlier F., Roche B., Autret S., Levine A., Seror S.J.;
RT "Sequencing of regions downstream of addA (98 degrees) and citC (289
RT degrees) in Bacillus subtlis.";
RL Microbiology 143:3305-3308(1997).
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RX MEDLINE-98044033; PubMed-9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertier M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brigneil S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaeser P., Goffeau A., Golligly E.J., Grand G.,
RA Guisepi G., Guy B.J., Haga K., Hatach J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Kobayashi Y., Koeltter P., Koningstein G., Krogh S., Kumano M.,
RA Kurioka K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Meisel C., Medigue C.,
RA Medina N., Mellado R.F., Mizuno M., Mostl D., Nakai S., Noack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogilawa A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portecelle D., Potwollik S., Prescott A.M.,
RA Prascan E., Puig C., Rochat E., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Schanlan E., Schleich S., Schroeter R., Scofield B.,
RA Seiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Solio B.,
RA Sotochun A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vanlier F., Vassarotti A.,
RA Violi A., Wamboldt R., Wedler E., Medler H., Wetzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RC -1- FUNCTION: NOT YET KNOWN. COULD BE INVOLVED IN PROTEOLYCAN
CC DegrADATION. BY CLEAVAGE OF ITS PEPTIDE BRIDGES.
CC -1- SUBCELLULAR LOCATION: CELL-WALL BOUND.
CC -1- PTM: PROCESSED INTO CMBP23 AND CMBP52.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U58681; AAC25926.1; -
DR EMBL: Y09476; CAAT0641.1; -
DR EMBL: Z99109; CAB12917.1; -
DR HSSP: Q45670; IDBI.
DR MEROPS: S08.004; -
DR SubtlisT: BG11846; WPA.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTLISIN.
DR PROSITE: PS00136; SUBTLISASE_ASP; FALSE_NEG.
DR PROSITE: PS00137; SUBTLISASE_HIS; 1.

```

DR	PROSITE: PS00138; SUBTILASE_SER; 1.
KM	Hydrolase: Serine protease; Cell wall; Zymogen; Signal;
FT	Complete proteome.
KM	Complete proteome.
FT	SIGNAL 1 31
FT	CHAIN 32 894
FT	CHAIN 32 ?
FT	PROBE 7 413
FT	CHAIN 414 894
FT	ACT_SITE 466 466
FT	ACT_SITE 497 497
FT	ACT_SITE 650 650
FT	CONFLICT 9 9
FT	CONFLICT 14 14
SO	SEQUENCE 894 AA; 96487 MW; 0F67C353E5F8DBC CRC64;
Query Match 8.4%; Score 262; DB 1; Length 894;	
Best Local Similarity 23.9%; Pred. No. 1.2e-11;	
Matches 134; Conservative 92; Mismatches 211; Indels 124; Gaps 28	
OY	49 FSKQXQTGAAP--LVESENVKLKGLKKLETPVA-----NKKLIHQFN---G 93
Db	245 FDVVSNGASSSYKTEDQKAIRNLRYDKALQSVSFLEIKEQADRLMMQLQGKTAG 304
OY	94 PLLE-----TKXLXTGAKLLIDYIPDAIYI----EYSG-----DYSXXXX 133
Db	305 ALLTENNTAAKEVQPT--KVLEFKVDKNSLSVHNEMKGFSAQAOKKDIDSNVRKKTL 362
OY	134 IEHVESVEPYELXXYYIDPOLFTFGASGXLYKAXALDPKOXKNREVQLGEIXAQ-XXXSND 192
Db	363 PNLXSFE--LPKDEKONAYTASAKRKYASAATLSKMNVAF---AEPOEYSLAND 416
OY	193 VXYITAKPEYKAMDVARGIV-ADV-AOSSGYLGOGO----IYAADTGLDGRNDS 246
Db	417 IY-----PYQMPLKNNGENGVKNADVKYEPAETLLSKRKLTDLIAVDTGDSTLAD-- 471
OY	247 MEAEFRGKITLAYA---LCRTNNANDTGHCHTHAVGSYLG---NCXTKNGMAPOANLVQ 300
Db	472 ----LKGVKRTDLGINPFGRRNNANNADDQHGCHTHAGIIAAGSDNGYSMTGLNAKKILIY 527
OY	301 SIMDSXGIGLGLPSNIQLTFESQASAGARIHTNSGCAAVNYATTDSRVNDYVRKNDAT 360
Db	528 KYLDSAG--SGDEQEIALGIKRYAADKGAVINLSIG----GGYSRLLEPALKYAADKNVL 581
OY	361 ILFAANGNEPNCGTISAPGTAKNALTTCATENLRSPFSYGADININHVAOFSSRGPTKGR 420
Db	582 IAAAGNGDEN--ALSTYPASSKYVMSVGAT-----NRMDVTADFSSYWGGL-- 625
OY	421 IKPDMAWPGTXILLARSRSIAPDSFWAMHDSKYAAMGTSMTPIYAGVNAOLREHFVN 480
Db	626 ---DISAGSDI-----PSLVPMGN-----VTYMSGSMATPYAAAAAGLLFPQNPL 670
OY	481 KQITFRKPSILKRA--LIAGADYG-----LGYP-----NGNOGWGVTYLDKSLN 522
Db	671 KTEVEDMLKKTAAADISFESYDGGEELYDDYGDIEIPKTPGVVMHSYGRILNMYKAVS 730
OY	523 VA-----YNNESSXLSTSOKAT 539
Db	731 AADDLQKVKKLESTOTAVRGS 751
RESULT 8	
Elya_BACSP STANDARD: PROT: 378 AA.	
ID	Elya_BACSP
AC	P20724:
DT	01-FEB-1991 (Rel. 17, Created)
DT	01-FEB-1991 (Rel. 17, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Alkaline elastase Yab precursor (EC 3.4.21.-).
GN	ALE
OS	Bacillus sp. (strain Yab).
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX	NCBI_TaxID=1409;
RN	[1]

```

RP      SEQUENCE FROM N.A.
RX      MEDLINE=89359181; PubMed=2670913;
RA      Kaneo R., Koyama N., Tsai Y.-C., Juang R.-Y., Yoda K., Yamasaki M.;
RT      "Molecular cloning of the structural gene for alkaline elastase Yab,
RL      a new subtilisin produced by an alkalophilic Bacillus strain.";
RN      J. Bacteriol. 171:5232-5236(1989).
RP      [2]
RA      SEQUENCE OF 111-164.
RT      Tsai Y.-C., Lin Y.-T., Li Y.-F., Yamasaki M., Tamura G.;
RL      "Characterization of an alkaline elastase from alkalophilic Bacillus
RT      ya-B.";
CC      Biochim. Biophys. Acta 883:439-447(1986).
CC      -|- FUNCTION: DIESTERIN EFFICIENTLY, HAS A SUBSTRATE PREFERENCE
CC      FOR ALA IN P POSITION.
CC      -|- SUBCELLULAR LOCATION: Secreted.
CC      -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL, M28537; AA87324.1; -.
DR      PIR, A33973; A33973.
DR      HSR, Q99405; 1MPT.
DR      MEROPS, S08.0PA; -.
DR      InterPro: IPR000209; Peptidase_S8.
DR      Pfam, PF00082; Peptidase_S8; 1.
DR      PRINTS; PRO0736; SUBTILISIN.
DR      PROSITE; PS00136; SUBTILASE_ASP; 1.
DR      PROSITE; PS00137; SUBTILASE_HIS; 1.
DR      PROSITE; PS00138; SUBTILASE_SER; 1.
KW      Hydrolase; Serine protease; zymogen; Signal.
FT      SIGNAL             1      27
FT      PROPEP            28     110
FT      CHAIN             111     378
FT      ACT_SITE          141     141
FT      ACT_SITE          171     171
FT      ACT_SITE          324     324
SQ      SEQUENCE          378 AA; 38793 MW; 5A8FB8CC6C62687D CnC64;

Query Match      8.3%; Score 257.5; DB 1; Length 378;
Best Local Similarity 26.7%; Pred. No. 8.4e-12;
Matches 115; Conservative 50; Mismatches 167; Indels 99; Gaps 20;

OY      73 LKKKLETPANNKLIHX-OFNGPILEETKQXLE--XTGAKILDYIPDAVYIEYGDVXS 129
DB      1 MKKKKGKIVAGALLIISAFSSIIQAAEEAKELIGFKEDVNSQ--FVDQIDGDEYS 58
OY      130 XXXXIIEHVESVEPYLPHYXIDPLEFTKGASKLVKAXALDTKQXNKEVOLRGLEIXIAQXXX 189
DB      59 ISSQAEDVEI--DLHERDFEIPVLSVELDEPVDALDPA----- 97
OY      190 SUDVYVYITAKPEYKMYNDVARGI--VKADVAQSSYGLYGQGIYAVADPTGIDGRNDSM 247
DB      98 --IAYIEDAEVTTMQTPVPGINRVPQIAQSR-GFTGTGVRVAVDLTGI-----SN 146
OY      248 HEAFGRKLTALYALGRTNANDNTNGHNHVGSV--LNGXGYNKMAQOANLVFOSIM-- 303
DB      147 HADLRIRGASGVPEEP-NISDGNHGVQVACTIALNNSIGVLGVAANVDLYGKVLGA 205
OY      304 DSXGGLGGLPSNLQTLFSQASAGARIHTNSWGAANGVAYTTDSNRVDVYRKNDMTILF 363
DB      206 SSGSGISGIAQILQ---WAANNNGHIANMSIGSAGSAGATMEQAVN---QATASGLVVA 258
OY      364 AAGNXPNGGTISAGCTAKNATVQATE--NLRPSFGYADINININVAQFSSGPKDKORI 421
DB      259 ASGNS--GAGNVGPFARYANMAVAQATDQNNRAFTFOYGAGL----- 299
OY      422 KPDVAAPGTXTILSARSSIALPDSFPAANDSKKAYVGGTSMATPIVAGVNA----- 471

```

```

Db 300 -DIVAPGV-----QSTVPNG-----YASFGTSMATPHVAGVAALVKQNPMSWS 345
Qy 472 -QJREHFVN 480
Db 346 NVQIRNH-LKN 355

RESULT 9
SUBN_BACNA STANDARD: PRT: 381 AA.
ID SUBN_BACNA
AC P35835;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Subtilisin NAT precursor (EC 3.4.21.62).
OS APRN.
GN Bacillus subtilis var. natto.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=86029;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NC-2;
RX MEDLINE=93113095; PubMed=1369081;
RA Nakamura T., Yamagata Y., Ichishima E.;
RT "Nucleotide sequence of the subtilisin NAT gene, aprN, of Bacillus
RT subtilis (natto).";
RL Biosci. Biotechnol. Biochem. 56:1869-1871(1992).
CC -1- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE,
CC IT CATALYZES THE HYDROLYSIS OF PROTEINS WITH BROAD SPECIFICITY
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in pI. Hydrolyzes, deplete amides.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF
CC SPOULATION, AND MANY MUTATIONS WHICH BLOCK SPOULATION AT EARLY
CC STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN
CC IS NOT NECESSARY FOR NORMAL SPOULATION.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D25319; BAA04989.1; -
DR EMBL: S51909; AAC60424.1; -
DR PIR: JH0778; JH0778.
DR HSSP: P07518; 1MEF.
DR MEROPS: S08.044.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PRO0723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILISIN_ASP; 1.
DR PROSITE: PS00137; SUBTILISIN_HIS; 1.
DR PROSITE: PS00138; SUBTILISIN_SER; 1.
KW Hydrolyase; Sporulation; Serine protease; Zymogen; Signal.
FT SIGNAL 1 23
FT PROPEP 24 106
FT CHAIN 107 381
FT ACT_SITE 138 138
FT ACT_SITE 170 170
FT ACT_SITE 337 327
SQ SEQUENCE 381 AA: 39507 MW: DAED4BI6ED1BA092 CRC64;

Query Match 8.2%; Score 255.5; DB 1; Length 381;
Best Local Similarity 25.9%; Pred. No. 1.2e-11;
Matches 124; Conservative 46; Mismatches 160; Indels 149; Gaps 19;
Qy 1 MRKKYFSLVLAAILSTVALXNPNRSGAXRXPDDFKGIQTITTDKXGFSKQXTGAAAF 60

```

```

Db 1 MRKKYFSLVLAAILSTVALXNPNRSGAXRXPDDFKGIQTITTDKXGFSKQXTGAAAF 60
Qy 61 LVESENVKLGK--LXKKLETPANNKLIHXOPNGPILEETKQXLEXTGATIDYIPDVA 118
Db 52 SAKKDVISEKGKVKQKRYVNA-----AAATDEKAVNELKDPESA 95
Qy 119 YIYVEGVDYXXXXXIEVESVEPYLPXYXIDPQLFTKASXLVAXAKDQKNKEVOL 178
Db 96 YVE-----DHAHEVAQSV-PY-----GISQ-IKAPL----- 122
Qy 179 RGIEIXAQQXXNDVXYITTAPEKYNADVAKIYKADVAQSGTGLGGQCIYAVADTGL 238
Db 123 -----HSGQITGSNKKVAVDSGI 141
Qy 239 DTGRDNDSEHAEFRGKITYALALGRTNANDTNGHGTAVGSV--LGNGXTNKGMAPQAN 296
Db 142 DSSHPLNV---RG--GASFPSEINPYODGSSHGTAVAGTIALNNSIGLVAPFAS 195
Qy 297 LVEQSIMDSXGGLGIPSNLTLEFSQAXSAGARIHTNSGAAVNGAVTTDSRNVDYVRK 356
Db 196 LVAVKVLDSGT--SGQYSWIINGIEMAINMDVIMSLGPGT---STALKTVVDKAVS 250
Qy 357 NDMTILFAGNEXPNG--GTISAPGTAKNAITVGA--TENVLPSPGAYADININVAQFS 412
Db 251 SGIVVAAAANGSGSGSTIVGYPAKPTSTIANGAVNSNGRASFSSVSGSEL----- 302
Qy 413 RGPTRKGRKPDVYAPGTXILSARSSLAIPDSSPMANHDSKYAYMGCTSMATPIYAGVA 471
Db 303 -----DYMAFGVSI-----QSTLPQGT-----YGVNGTSMATPHVAGAA 338

RESULT 10
AQL1_THEAQ STANDARD: PRT: 513 AA.
ID AQL1_THEAQ
AC P08594;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aqualysin I precursor (EC 3.4.21.-).
GN PSTI.
OS Thermus aquaticus.
OC Bacteria; Thermus/Delnoococcus group; Delnoococi; Thermates;
OC Thermaceae; Thermus.
OX NCBI_Taxid=271;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 15-23.
RC STRAIN=YT1.
RX MEDLINE=90216674; PubMed=2182621;
RA Terada I., Kwon S.-T., Miyata T.;
RT "Unique precursor structure of an extracellular protease, aqualysin
RT I, with NH2- and COOH-terminal pro-sequences and its processing in
RT Escherichia coli.";
RL J. Biol. Chem. 265:6576-6581(1990).
RN [2]
RP SEQUENCE OF 75-442 FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=YT1.
RX MEDLINE=88225062; PubMed=2286255;
RA Kwon S.-T., Terada I., Matsuzawa H., Ohta T.;
RT "Nucleotide sequence of the gene for aqualysin I (a thermophilic
RT alkaline serine protease) of Thermus aquaticus YT-1 and
RT characteristics of the deduced primary structure of the enzyme.";
RL Eur. J. Biochem. 173:491-497(1988).
RN [3]
RP SEQUENCE OF 128-170.
RX MEDLINE=88151937; PubMed=3162211;
RA Matsuzawa H., Tokugawa K., Hamockl M., Mizoguchi M., Taguchi H.,
RA Terada I., Kwon S.-T., Ohta T.;
RT "Purification and characterization of aqualysin I (a thermophilic
RT alkaline serine protease) produced by Thermus aquaticus YT-1.";
RL Eur. J. Biochem. 171:441-447(1988).
CC -1- FUNCTION: AQUALYSIN I IS A THERMOPHILIC ALKALINE SERINE PROTEASE.
CC THE OPTIMAL TEMPERATURE FOR ITS CASEINOLYTIC ACTIVITY IS 80

```


[8]
RN ACTIVE SITE.
RX MEDLINE=69104413; PubMed=5249818;
RA Markland F.S., Shaw E., Smith E.L.;
RT "Identification of histidine 64 in the active site of subtilisin";
RL Proc. Natl. Acad. Sci. U.S.A. 61:1440-1447(1968).
CC -1- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE,
CC IT CATALYZES THE HYDROLYSIS OF PROTEINS AND PEPTIDE AMIDES.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in P1. Hydrolyzes peptide amides.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- BIOTECNOLOGY: Used as a detergent protease. Sold under the name
CC Alcalase by Novozymes.
CC -1- MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF
CC SPOROGATION AND MANY MUTATIONS WHICH BLOCK SPOROGATION AT EARLY
CC STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN
CC IS NOT NECESSARY FOR NORMAL SPOROGATION.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K02496; AAB05345.1; -
DR PIR: A00970; SUBSN.
DR PDB: 1S01; 15-OCT-90.
DR PDB: 1S02; 15-JAN-92.
DR PDB: 1SBH; 07-DEC-95.
DR PDB: 1SB1; 07-DEC-95.
DR PDB: 1SBN; 31-JAN-94.
DR PDB: 1SRT; 31-MAY-84.
DR PDB: 2SRT; 31-MAY-84.
DR PDB: 1STB; 31-OCT-93.
DR PDB: 2STC; 15-APR-93.
DR PDB: 3STC; 31-JAN-94.
DR PDB: 5STC; 31-JAN-94.
DR PDB: 2SNT; 15-JAN-93.
DR PDB: 1SEB; 15-OCT-95.
DR PDB: 2ST1; 15-JUL-91.
DR PDB: 1ST2; 15-JUL-91.
DR PDB: 1SUA; 14-JAN-98.
DR PDB: 1SUB; 31-JAN-94.
DR PDB: 1SUC; 31-JAN-94.
DR PDB: 1SUD; 31-JAN-94.
DR PDB: 1SUE; 14-OCT-98.
DR PDB: 1SUF; 14-NOV-95.
DR PDB: 1AK9; 12-NOV-97.
DR PDB: 1AUG; 31-DEC-97.
DR PDB: 1AON; 14-JAN-96.
DR PDB: 1YJA; 11-JUL-96.
DR PDB: 1YJB; 11-JUL-96.
DR PDB: 1YJC; 11-JUL-96.
DR PDB: 1A20; 29-APR-98.
DR MEMOS: S08 034.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Sporulation; Serine protease; Zymogen; Signal;
KW 3D-structure.
FT SIGNAL 1 32 POTENTIAL.
FT PROPEP 33 107
FT CHAIN 108 382 SUBTILISIN BPN'.
FT ACT_SITE 139 139 CHARGE RELAY SYSTEM.
FT ACT_SITE 171 171 CHARGE RELAY SYSTEM.
FT ACT_SITE 328 328 CHARGE RELAY SYSTEM.

FT HELIX 113 117
FT TURN 118 119
FT HELIX 120 125
FT TURN 126 127
FT TURN 131 132
FT STRAND 134 135
FT TURN 144 145
FT TURN 147 148
FT STRAND 151 156
FT TURN 159 160
FT TURN 164 165
FT HELIX 171 180
FT TURN 193 194
FT STRAND 196 201
FT TURN 205 206
FT HELIX 211 223
FT TURN 224 225
FT STRAND 228 231
FT STRAND 235 235
FT HELIX 240 251
FT TURN 252 253
FT STRAND 255 259
FT TURN 267 268
FT STRAND 274 274
FT TURN 275 277
FT TURN 279 280
FT STRAND 282 287
FT TURN 289 290
FT STRAND 293 293
FT TURN 295 296
FT TURN 301 302
FT STRAND 305 308
FT TURN 312 316
FT TURN 317 319
FT STRAND 320 324
FT HELIX 327 344
FT TURN 346 347
FT TURN 350 359
FT HELIX 350 359
FT STRAND 362 362
FT STRAND 367 370
FT TURN 371 372
FT STRAND 374 374
FT HELIX 377 381
SQ SEQUENCE 382 AA; 39181 MW; ED987DAFAJ7B8335 CRC64;

Query Match 8.2%; Score 253; DB 1; Length 382;
Best Local Similarity 24.0%; Pred. No. 1.8e-11;
Matches 120; Conservative 43; Mismatches 148; Indels 188; Gaps 17;

QY 1 MRKKVFLSVLSAAIIISTVALKNPMSGAKRFDLPKGIOTTTXKGEKXOTGAAP 60
DB 1 MRKKVWISLPLALILFTTMAFGSTSAQAQ-----KNGEKKIYGFQOTMTMSAA 54

QY 61 LVSEENVKLKGLAKKLETPVANKKLHIQFNGPILEETKOKLXETGAKIITYPYAYI 120
DB 55 -----KKKQVISEKGVQKQFK--YDAAASATLNKAAVKELKQPSVAVY 98

QY 121 VEYEGVXSXXXXIEHV-----ESVEPYLPXYXIDPLFTKGASLYVAKALDTKQXNKE 175
DB 99 EE-----DHVAHAQAQSV-py-----GVSO-IKAPAL----- 123

QY 176 VQURGLEIXIAQXXSNDVXYITAPKPYKVMNDVARGIVKADVAQSVGLGQGIIVAVAD 235
DB 124 -----HSQGYTSNKKVAVID 139

QY 236 TGLDTRGNDSSMHEAFRGKITATYALGRTNNANDTGHGTHVAGSV--LGNGXTNKKMAP 293
DB 140 SGID-----SSHPLKVAAGASVWPSETNPFDNNSHGTHVAGTVAALNNSIGVLGVA 193

QY 294 QANLVPOSITMDSXGLGLGSPNLQTLFSQAXSAGARIHTNSMGAAVNCVATYTTDSBRND-- 351
DB 194 SASTLVAKVLAGDG-----SGQYSLWLNQIEWALANNDVY 229

```

QY 352 -----D-YRKNMUTLIFAGNEXPNG--GTISAPGTAKNAITVGATEN 392
D 230 NMSLGGPSSAALKAAYKAVASGVVVAAGNEGTSSSSSTVGYPKYPVLAVGA--- 286
QY 393 LRPSGSGADININHAQSSSRGPKDGRIKPDVMAFGYXILSARSLAPDSSFMANHRSK 452
D 287 -----VDSSNQRASFSVGP-----ELDVMAFGVSIQSTLPG-----NK 320
QY 453 YAYMGTSMATPIVAGNVA 471
D 321 YGAYNGTSMASPHVAGAA 339

RESULT 12
SUBT_BACSA
ID SUBT_BACSA STANDARD: PRT: 381 AA.
AC P00783;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Subtilisin amylosacchariticus precursor (EC 3.4.21.62).
GN APR.
OS Bacillus subtilis var. amylosacchariticus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1483;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE=89008194; PubMed=3139650;
RA Yoshimoto T., Oyama H., Honda T., Tone H., Takeshita T.,
RA Kamlyama T., Tsuru D.;
RT "Cloning and expression of subtilisin amylosacchariticus gene.";
RL J. Biochem. 103:1060-1065(1988).
RN 12;
RP PARTIAL SEQUENCE.
RX MEDLINE=72266687; PubMed=4560201;
RA Markland F.S., Kurihara M., Smith E.L.;
RT "Subtilisin Amylosacchariticus. II. Isolation and sequence of the
RT tryptic and cyanogen bromide peptides.";
RL J. Biol. Chem. 247:5602-5618(1972).
RN 13;
RP SEQUENCE OF 107-381.
RX MEDLINE=72266688; PubMed=5055784;
RA Kurihara M., Markland F.S., Smith E.L.;
RT "Subtilisin Amylosacchariticus. 3. Isolation and sequence of the
RT chymotryptic peptides and the complete amino acid sequence.";
RL J. Biol. Chem. 247:5619-5631(1972).
CC -1- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE,
CC IT CATALYZES THE HYDROLYSIS OF PROTEINS AND PEPTIDE AMIDES.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in pI. Hydrolyzes peptide amides.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF
CC SPOROGATION, AND MANY MUTATIONS WHICH BLOCK SPOROGATION AT EARLY
CC STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN
CC IS NOT NECESSARY FOR NORMAL SPOROGATION.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, D00264; BAA00186.1; -.
DR PIR, A00971; SUBSS.
DR PIR, A41448; A41448.
DR HSSP, P04189; ISCJ.
DR MEROPS, S08.042; -.
DR Interpro, IPR000209; Peptidase_S8.

```

```

DR Pfam, PF00082; Peptidase_S8; 1.
DR PRINTS, PR00723; SUBTILISIN.
DR PROSITE, PS00136; SUBTILISIN_ASP; 1.
DR PROSITE, PS00137; SUBTILISIN_HIS; 1.
DR PROSITE, PS00138; SUBTILISIN_SER; 1.
KM Hydrolyase; Sporulation; Serine protease; Zymogen; signal.
FT SIGNAL 30
FT PROPEP 31 106
FT CHAIN 107 381
FT ACT_SITE 138 138
FT ACT_SITE 170 170
FT ACT_SITE 327 327
FT CONFLICT 191 191
FT CONFLICT 365 365
SQ SEQUENCE 381 AA; 39467 MW; 2251BADE22B4824F CRC64;

Query Match 8.1%; Score 251.5; DB 1; Length 381;
Best Local Similarity 26.0%; Pred. No. 2.3e-11;
Matches 124; Conservative 49; Mismatches 159; Indels 145; Gaps 19;

QY 1 MRKKVFLSVLAAILSTVALXNPSAGXARFDFKGIQTTDXGFSKQXGTGAFAF 60
D 1 MRKKLWISLFLALTLITMAFNSMAQAGKSSTEKRYI-----VGFK---QTMSSAMS 51
QY 61 LVESENVKLXKG-LKKRLTVPANNKLHIXQFNPGPILEETKQXLEXTGAKILDYIPDYA 118
D 52 SAKKKDVISEKRGKVKQKQFYVNA-----AAATLDEKAVKELKKDP5VA 95
QY 119 YIVEEGVAXKXXXIENVESEVEPLPYXITDPLQFTGASXLKAKALDTRQAKKEVQL 178
D 96 YVEE-----DHIAHEYAQSV-PY-----GISQ-IKAPAL----- 122
QY 179 RGIEIXAQXXXNDVXYITAKPEKVMNDVARGIYKADVAQSSYGLYQOGQIVAVADTGL 238
D 123 -----HSQCYGSNKKVAVINDSCI 141
QY 239 DTGRDSSMHEAFRGKITALYALGRTNANDTNGHTHVAGSV-LGNGXTNKGAPQAN 296
D 142 DSHPLDLNV-----RG--CASFPVSETPNYQDSSSHGTAGTIALNNSICVLAVPSAS 195
QY 297 LVFOSIMDSXGGLGLPNSLQTLFSQASAGARIHTNMGCAVAVNGAYITDSRNVDDYRK 356
D 196 LYAVKVLDTSTG--SGQYSWIINGIEMWATSNMMDYINMSLGP---SGSTAKTAYVDKAVS 250
QY 357 NDMTILFAGNEXPNG--GTISAPGTAKNAITVGATENLRPSFGSYADININHAQSSRG 414
D 251 SGIVVAALAGNMGSSSTYGVPAKYPSTIYAVGA-----VSSNQRASFSFSG 299
QY 415 PTKDGRIKPDVMAFGYXILSARSLAPDSSFMANHRSKRYAYMGTSMATPIVAGNVA 471
D 300 S-----ELDVMAFGVSI-----QSTLPGCT-----YGAYNGTSMATPHVAGAA 338

RESULT 13
ELIA_BACAO
ID ELIA_BACAO STANDARD: PRT: 380 AA.
AC P27693;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alkaline protease precursor (EC 3.4.21.-).
OS Bacillus alcalophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1445;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=PB92;
RX MEDLINE=91282483; PubMed=2059048;
RA van der Laan J.C., Gerritse G., Mulleners L.J.M., van der Hoek R.A.,
RA Quax W.J.;
RT "Cloning, characterization, and multiple chromosomal integration of a
RT Bacillus alkaline protease gene.";
RL Appl. Environ. Microbiol. 57:901-909(1991).

```

```

RN [2]
RA P-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
RC STRAIN-PB92;
RD MEDLINE=92390330; PubMed=1518768;
RE van der Laan J.C., Teplovskoy A.V., Kelders H., Kalk R.H., Misset O.,
RF Molkeniers L.J.M., Dijkstra B.W.;
RG "crystal structure of the high-alkaline serine protease PB92 from
RH Bacillus alcalophilus."
RI Protein Eng. 5:405-411(1992).
RL [3]
RA X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RC MEDLINE=93078250; PubMed=1447775;
RD Sobek H., Hecht H.-J., Aehle W., Schomburg D.;
RE "x-ray structure determination and comparison of two crystal forms of
RF a variant (Asn114Arg) of the alkaline protease from Bacillus
RG alcalophilus refined at 1.85-A resolution."
RH J.Mol. Biol. 228:108-117(1992).
RN [4]
RA STRUCTURE BY NMR OF 112-380.
RC STRAIN-PB92;
RD MEDLINE=92727237; PubMed=9115441;
RE Marini J.R., Mulder F.A., Kariml-Nejad Y., van der Zwan J.,
RF Marini M., Schipper D., Boelens R.;
RG "the solution structure of serine protease PB92 from Bacillus
RH alcalophilus presents a rigid fold with a flexible substrate-binding
RI site."
RJ Structure 5:521-532(1997).
CC -1 SUBCELLULAR LOCATION: Secreted.
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M65086; AAA22212.1; -
DR EMBL: A13738; CA01128.1; -
DR PIR: A49778; A49778.
DR PDB: 1AH2; 15-APR-98.
DR MEROPS: S08.038.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Serine protease; Tyrosinase; Signal; 3D-structure.
FT SIGNAL 1 27
FT PROPEP 28 112 POTENTIAL.
FT CHAIN 113 380 ALKALINE PROTEASE.
FT ACT_SITE 143 143 CHARGE RELAY SYSTEM.
FT ACT_SITE 173 173 CHARGE RELAY SYSTEM.
FT ACT_SITE 326 326 CHARGE RELAY SYSTEM.
SQ SPOUNCE 380 AA; 38853 MW; 539EA727171B6682C CRC64;

Query Match 8.1%; Score 250.5; DB 1; Length 380;
Best Local Similarity 24.4%; Pred. No.2.8e-11;
Matches 123; Conservative 57; Mismatches 148; Indels 177; Gaps 22;

Oy 4 KYKFVLSVAIAAIIISTVALKNPSGAKARFPLDKGIQTITDDXSGSKQXTGAATVE 63
Db 2 KRFLGIVASTALLISAVSSSIASAEAAEKY-----LIGNED----- 42

Oy 64 SEWVKXKGKKIKKEYPANNNKHITQFNPGPILESFTQKLEHXGAKILDYIPDAIYVER 123
Db 43 -----ENVEPFQYEANDAEVAI-----LSEEEVEIE-----LHFF 75

Oy 124 EG-DVAXSXKXXIXEHVESVEPYLPFXKIIDPOLFTFGASXLVKAAALDTOKXNKVEQLRGIE 182
Db 76 ETIPVLVSLESPDEADALE-----LDPA----- 98

```

```

Oy 183 XIAQXXXSXNDXYITAKPEKKV-NDVARGIKADV--AOSSTYLQGOQIYAADTLDI 240
Db 99 -----1SYLEEDAEVTTMAQGVPMGDISRQAPAAINRGLTSSGVKAVYDGTIST 148
Oy 241 GRNDSMHAEAFRCGKITLALYALGTNNADNTDGHCTHVASGV--LGGXTXNKGMAPONLV 298
Db 149 -----HPDLININGGASFLVGPGR-STQDGNCGHGTHTVAGTIALNNSIGVLGAPANMEL 200
Oy 299 FQSIDSKXGGLGGLPSMLQTLFSQAXSAGARLHTNSGA-----AVNGATTTDSRN 349
Db 201 AYKVLGASG--SGSVSTIAQGLEMGNGNNHVMNLISGSPSPATLEQAVNSA---TSRG 235
Oy 350 VDDYRKNDMTITLFAAGNEKPNCGTISAPGAKNAITVGATE--NLRPSPGSYADNIHV 407
Db 256 V-----LVVAASGN--GAGSTISPARYAAAMVGGATDDNNRRASFQOYAGL--- 301
Oy 408 AOPSSNGPTKDGRIKRPDYMAPGTXILISARRSLAPDSSFMANHDHKKAYVAGTSMATPIVA 467
Db 302 -----DIVAPGVNOVSTPG-----SYVASLNGTSMATPHVA 333
Oy 468 GNVV-----OLRHEHFEKN 480
Db 334 GAALVYKQKNPSMSNQIRNH-LKN 357

RESULT 14
ELIXA_BACCS
ID ELIXA_BACCS STANDARD: PRF: 380 AA.
AC P41362;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alkaline protease precursor (Ec 3.4.21.-).
OS Bacillus clausii.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=79880;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;
RX MEDLINE=93043753; PubMed=1368952;
RA Takami H., Kobayashi T., Kobayashi M., Yamamoto M., Nakamura S.,
RA Aono R., Horikoshi K.;
RT "Molecular cloning, nucleotide sequence, and expression of the
RT structural gene for alkaline serine protease from alkaliphilic
RT Bacillus sp. 221."
RL Biosci. Biotechnol. Biochem. 56:1455-1460(1992).
RN [2]
RP SEQUENCE OF 112-129.
RC STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;
RA Horikoshi K.;
RL (in) Horikoshi K. (eds.);
RL Microorganisms in alkaline environments, pp.187-194, VCH,
RL Weinheim (1991).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.jsb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DB EMBL, S48754, AAC60420.1, -
DB EMBL, D13157, BAA02442.1, -
DB EMBL, A26817, CA01836.1, -
DB EMBL, A22350, CA01611.1, -
DB HSSP, P29600, ICGI.
DB MEROPS, S08.103, -
DB InterPro, IPR000209, Peptidase_S8.
DB Pfam, PF00082, Peptidase_S8, 1.
DB

```

```

DR PRINTS: PRO0723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE ASP; 1.
DR PROSITE: PS00137; SUBTILASE HIS; 1.
DR PROSITE: PS00138; SUBTILASE SER; 1.
KW Hydrolyase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 27
FT PROPEP 28 111
FT CHAIN 112 380
FT ACT_SITE 143 143
FT ACT_SITE 173 173
FT ACT_SITE 326 326
SO SEQUENCE 380 AA; 38826 MW; 5F73ABC68D586831 CRC64;

Query Match      8.1%; Score 250.5; DB 1; Length 380;
Best Local Similarity 24.4%; Pred. No. 2.8e-11;
Matches 123; Conservative 57; Mismatches 148; Indels 177; Gaps 22;

QY 4 KKVFLSVLSAAALISTVALKNPSAGAKXFDLDFKGIOTTDXXGFSKXQGTGAAPLVE 63
DB 2 KKPLGKIYASTALLISVAFSSSIASNAEAEKRY-----LIGFNEQ----- 42
QY 64 SENVLXGKGLKKTETVANKNLHXQNGPILEETKQXLETKGAKIIDYIPDAVIYEX 123
DB 43 -----EAVSEFEQVANDQVAL-----LSEEEVEITE-----LHFE 75
QY 124 EG-DYKXKXKXKXIEHVESEPYLPXYXIDPOLFTKGASLXVKAXALDRQXNKQVQLRGIE 182
DB 76 ETIPLVSLVELSPEDVDAL-----LDPA----- 98
QY 183 XIAQKXKXNDYXITAKEPEYKX-NDVARGIYKADY-AOSSYGLYGOGIYAVADTGLDT 240
DB 99 -----ISYIEDAEVTTTMAQSVPMGSRVQAPAHNNGLSGKXAVYLDLTGIST 148
QY 241 GRNDSMEHFAFGKITIYALGRTNANDTNGCHGVHVASV--LGNGXTNKGMADQANLY 298
DB 149 -----HPDLNRGGAFFYFGEF-STODGNCHGVHAGTIALALNNSIGVLCVAFSAELY 200
QY 299 FOSIMDSXGGLGGLPSNLQTLFSQAXSAGARIHTNSMGA-----AVNGAYTTDSRN 349
DB 201 AVKVLGASG--SGSVSIAQGLEMGAGNNGMHVANAISLSPSPSATLEQAVNSA---TSRG 255
QY 350 VDDYRKNDMTLFFAAGXENPGRTISAPGTAKNAITVGAP--NLRRPSFGSYADININIV 407
DB 256 V-----LVVAASGNS--GAGSISTPYARANAAGATDQNNRASFSSQYAGL--- 301
QY 408 AOFSSRGPTKDGRIKPDYKAPGTXLISARSSLAPDSFWANHDSKYAYMGTSMAPTIYA 467
DB 302 -----DIVAPGVNVQSTYTPG-----STYASLNGTSMATPHYA 333
QY 468 GNVA-----QLRHFVKX 480
DB 334 GAALVYKOKNPSWNSVQIRNH-LKN 357

RESULT 15
SUBT_BACLI STANDARD; PRT; 379 AA.
ID SUBT_BACLI
AC P00780;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Subtilisin Carlsberg precursor (EC 3.4.21.62).
GN APR.
OS Bacillus licheniformis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1402;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIB 6816;
RX MEDLINE=86093688; PubMed=3001653;
RA Jacobs M., Eliasson M., Uhlen M., Flock J.-I.;
RT "Cloning, sequencing and expression of subtilisin Carlsberg from
  Bacillus licheniformis.";

```

```

RL Nucleic Acids Res. 13:8913-8926(1985).
RN [2]
RP SEQUENCE OF 106-379.
RX MEDLINE=68234702; PubMed=4967581;
RA Smith E.L., Delange R.J., Evans W.H., Landon M., Martland F.S.;
RT "Subtilisin Carlsberg. V. The complete sequence; comparison with
  subtilisin BPN'; evolutionary relationships.";
RL J. Biol. Chem. 243:2184-2191(1968).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF VARIANT WITH SELENOCYST-325.
RX MEDLINE=93291170; PubMed=8512925;
RA Syed R., Wu Z.P., Hogle J.M., Hillvert D.;
RT "Crystal structure of selenosubtilisin at 2.0-A resolution.";
RN Biochemistry 32:6157-6164(1993).
RP [4]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 106-379.
RX MEDLINE=98087517; PubMed=9425066;
RA Stoll V.S., Eger B.T., Hynes R.C., Martichonok V., Jones J.B.,
  Pal E.F.;
RT "Differences in binding modes of enantiomers of L-acetamido boronic
  acid based protease inhibitors: crystal structures of gamma-
  chymotrypsin and subtilisin Carlsberg complexes.";
RL Biochemistry 37:451-462(1998).
CC -1- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE,
  IF CATALYZES THE HYDROLYSIS OF PROTEINS AND PEPTIDE AMIDES.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
  for peptide bonds, and a preference for a large uncharged residue
  in pl. Hydrolyzes peptide amides.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- BIOTECHNOLOGY: Used as a detergent protease. Sold under the name
  Alcalase by Novozymes.
CC -1- MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF
  SPOULATION, AND MANY MUTATIONS WHICH BLOCK SPOULATION AT EARLY
  STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN
  IS NOT NECESSARY FOR NORMAL SPOULATION.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X03341; CAB56500.1; -.
DR PIR: A00968; SUBSC.
DR PIR: A24111; SUBSCL.
DR PDB: 1CSE; 15-OCT-89.
DR PDB: 1SCA; 31-JAN-94.
DR PDB: 1SCB; 31-JAN-94.
DR PDB: 1SCD; 31-JAN-94.
DR PDB: 1SBC; 15-JAN-95.
DR PDB: 2SEC; 15-JAN-95.
DR PDB: 1SCN; 31-AUG-94.
DR PDB: 1AF4; 16-JUN-97.
DR PDB: 1SEL; 31-OCT-93.
DR PDB: 1VSB; 18-MAR-98.
DR PDB: 3VSB; 25-MAR-98.
DR PDB: 1AV7; 01-APR-98.
DR PDB: 1BE6; 14-OCT-98.
DR PDB: 1BE8; 13-JAN-99.
DR PDB: 1BFK; 18-NOV-98.
DR PDB: 1BFU; 18-NOV-98.
DR MEROPS: S08.001; -.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PRO0723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE ASP; 1.
DR PROSITE: PS00137; SUBTILASE HIS; 1.
DR PROSITE: PS00138; SUBTILASE SER; 1.
KW Hydrolyase; Sporulation; Serine protease; Zymogen; Signal;

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:35:29 ; Search time 60.4527 Seconds
(without alignments)
2177.969 Million cell updates/sec

Title: US-09-920-954-1

Perfect score: 3101
Sequence: 1 MRKKVFELSVSAALISTV.....EVQATNPVGPQKSLAIVN 639

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mmc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.todent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.rvirus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2155	69.5	434	2	Q9AQR3 Bacillus sp
2	2155	69.5	434	2	Q9AQR3 Bacillus sp
3	2155	69.5	434	2	Q9AQR3 Bacillus sp
4	1952.5	63.0	433	2	Q9AQR1 Bacillus sp
5	1948.5	62.8	433	2	Q9AQR1 Bacillus sp
6	1941.5	62.6	433	2	Q9AQR2 Bacillus sp
7	493	15.9	1825	5	Q8TRW1 dictyostell
8	478	15.4	1702	5	Q9GTW7 dictyostell
9	412	13.3	561	16	Q8RBJ2 thermomae
10	408	13.2	654	17	Q8UBC9 pyrococ
11	391	12.6	1239	16	Q9FZ4 streptomyc
12	368.5	11.9	1253	16	Q9FC6 streptomyc
13	356	11.3	1102	2	P95684 streptomyc
14	350.5	11.3	444	16	Q9KRJ7 bacillus ha
15	337	10.9	1245	16	Q9R54 streptomyc
16	336.5	10.9	412	2	Q9AER6 thermomae

17	336.5	10.9	412	16	Q8RC68 thermomae
18	321.5	10.4	1220	16	Q9L0A0 streptomyc
19	319.5	10.3	442	2	Q31788 bacillus su
20	315	10.2	824	2	Q45464 bacillus su
21	310	10.0	891	1	Q93635 thermococ
22	301.5	9.7	1398	1	Q9P9L1 pyrococ
23	295.5	9.5	431	2	Q9S3L6 bacillus sp
24	281.5	9.1	434	2	Q54327 bacillus sp
25	280	9.0	1135	1	Q9P9D1 uncultured
26	279.5	9.0	419	2	Q45681 bacillus su
27	276.5	8.9	799	16	Q9KEM1 thermomae
28	270.5	8.7	621	2	P97097 streptomyc
29	270	8.7	621	2	Q9F486 alteromon
30	270	8.7	621	2	Q53401 alteromon
31	269	8.7	379	2	Q66153 bacillus sp
32	268.5	8.7	715	2	P70765 alteromon
33	260.5	8.4	1345	1	Q54437 staphylothe
34	257.5	8.3	403	2	Q45463 bacillus sp
35	255	8.2	692	2	Q9EXK0 bacillus ps
36	253	8.2	1098	16	Q9L1Z8 streptomyc
37	253	8.2	629	2	Q93RG8 alteromon
38	250.5	8.1	374	2	Q9FP943 bacillus 11
39	250.5	8.1	374	2	Q9FP942 bacillus 11
40	250.5	8.1	378	2	Q45466 bacillus sp
41	250.5	8.1	601	2	Q46540 bacteroides
42	250	8.1	379	2	Q45300 bacillus 11
43	250	8.1	382	2	Q45522 bacillus sp
44	249.5	8.0	374	2	Q9P941 bacillus 11
45	249.5	8.0	757	16	Q9K6G6 bacillus ha

ALIGNMENTS

RESULT 1
Q9AQR3 PRELIMINARY; PRT; 434 AA.
AC Q9AQR3;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Protease (Fragment).
GN PROA.
OS Bacillus sp. 9860.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacilliales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=133778;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-9860.
RX MEDLINE-20568675; PubMed-11118284;
RA Sasaki K., Okuda M., Hatada T., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046403; BAB21266.1; -.
DR HSSP; P00782; 1SUP.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 2.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
FT NON_TER 1
FT NON_TER 434
SQ SEQUENCE 434 AA; 45311 MW; AFC9F78EB143527E CRC64;
Query Match 69.5%; Score 2155; DB 2; Length 434;
Best Local Similarity 96.3%; Pred. No 2, 1e-122;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
OY 206 NDVARGIVADVAQSSYGLYGCGQIVAVADTGLDTRGNDNSMHEAFKCKITATLALGRTN 265

```
|||||
Db 1 NDVARGIKADVAOSSYGLYGOGQIVAAVDGLDGRNDSSMHEAFRGKITALVALGRTN 60
Qy 266 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVFQSIIMDSXGGLGCLPSNLQTLFSQAXS 325
Db 61 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVFQSIIMDSXGGLGCLPSNLQTLFSQAXS 120
Qy 326 AGARITHMSWGAANVGAYTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
Db 121 AGARITHMSWGAANVGAYTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
Qy 386 TVGATENIRPSPGSIYADININHYAOPSSRGPTKDGRIKPDVMAPGTXILSARSLAPDSSF 445
Db 181 TVGATENIRPSPGSIYADININHYAOPSSRGPTKDGRIKPDVMAPGTXILSARSLAPDSSF 240
Qy 446 WANHDSKYAYMGTSMAPIVAGNVAOLREHFVKNRGITTPKPSLKALILAGADXGLGY 505
Db 241 WANHDSKYAYMGTSMAPIVAGNVAOLREHFVKNRGITTPKPSLKALILAGADVGLGY 300
Qy 506 PNGNGMGWRVTLDKSLNVAAYNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 565
Db 301 PNGNGMGWRVTLDKSLNVAAYNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 360
Qy 566 SVTLVNDLDVLTAPNGTYVGNDFXPPXXMMDCGRNVENFIIXPQSGTYTIEVQAYN 625
Db 361 SVTLVNDLDVLTAPNGTYVGNDFXPPXXMMDCGRNVENFIIXPQSGTYTIEVQAYN 420
Qy 626 VPVGPQXFSLAIVN 639
Db 421 VPVGPQXFSLAIVN 434
```

RESULT 2

```
093UV9 PRELIMINARY: PRT: 434 AA.
ID 093UV9.
AC 093UV9.
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Protease (Fragment).
GN PROE.
OS Bacillus sp. KSM-KP43.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=109322;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KP43.
RA Itoh S., Saeki K.;
RT "new protease.";
RL Submitted (NCV-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB051423; BAB55674.1; -
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 434
SQ SEQUENCE 434 AA: 45302 MM: 0B08418191853CA3 CRC64;
```

Query Match 69.5%; Score 2155; DB 2; Length 434;
Best Local Similarity 96.3%; Pred. No. 2,1e-122;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy 206 NDVARGIKADVAOSSYGLYGOGQIVAAVDGLDGRNDSSMHEAFRGKITALVALGRTN 265
Db 1 NDVARGIKADVAOSSYGLYGOGQIVAAVDGLDGRNDSSMHEAFRGKITALVALGRTN 60
Qy 266 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVFQSIIMDSXGGLGCLPSNLQTLFSQAXS 325
Db 61 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVFQSIIMDSXGGLGCLPSNLQTLFSQAXS 120
Qy 326 AGARITHMSWGAANVGAYTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
```

```
|||||
Db 121 AGARITHMSWGAANVGAYTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
Qy 386 TVGATENIRPSPGSIYADININHYAOPSSRGPTKDGRIKPDVMAPGTXILSARSLAPDSSF 445
Db 181 TVGATENIRPSPGSIYADININHYAOPSSRGPTKDGRIKPDVMAPGTXILSARSLAPDSSF 240
Qy 446 WANHDSKYAYMGTSMAPIVAGNVAOLREHFVKNRGITTPKPSLKALILAGADXGLGY 505
Db 241 WANHDSKYAYMGTSMAPIVAGNVAOLREHFVKNRGITTPKPSLKALILAGADIGLY 300
Qy 506 PNGNGMGWRVTLDKSLNVAAYNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 565
Db 301 PNGNGMGWRVTLDKSLNVAAYNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 360
Qy 566 SVTLVNDLDVLTAPNGTYVGNDFXPPXXMMDCGRNVENFIIXPQSGTYTIEVQAYN 625
Db 361 SVTLVNDLDVLTAPNGTYVGNDFXPPXXMMDCGRNVENFIIXPQSGTYTIEVQAYN 420
Qy 626 VPVGPQXFSLAIVN 639
Db 421 VPVGPQXFSLAIVN 434
```

RESULT 3

```
09AOR0 PRELIMINARY: PRT: 434 AA.
ID 09AOR0.
AC 09AOR0.
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Protease (Fragment).
GN PROE.
OS Bacillus sp. NVL.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=133781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NVL.
RA MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL: AB046406; BAB21269.1; -.
DR HSSP: P00782; 1SOP.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 2.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
FT NON_TER 1
FT NON_TER 434
SQ SEQUENCE 434 AA: 45294 MM: 83517EDDB74125D2 CRC64;
```

Query Match 67.1%; Score 2082; DB 2; Length 434;
Best Local Similarity 91.5%; Pred. No. 5.3e-118;
Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

```
Qy 206 NDVARGIKADVAOSSYGLYGOGQIVAAVDGLDGRNDSSMHEAFRGKITALVALGRTN 265
Db 1 NDVARGIKADVAOSSYGLYGOGQIVAAVDGLDGRNDSSMHEAFRGKITALVALGRTN 60
Qy 266 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVFQSIIMDSXGGLGCLPSNLQTLFSQAXS 325
Db 61 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVFQSVIMDSXGGLGCLPSNVSTLFSQAXS 120
Qy 326 AGARITHMSWGAANVGAYTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
Db 121 AGARITHMSWGAANVGAYTTDSRNVDYVRKNDMAVLFAAGNEXPNGGTISAPGTAKNAI 180
```

```
QY 386 TVGATENLRSPFGSYADNINNHVAQFSSRGPTKDGRIKPDVMAPTXTLLSARSSLAPDSSF 445
|||||
DB 181 TVGATENLRSPFGSYADNINNHVAQFSSRGPTKDGRIKPDVMAPTXTLLSARSSLAPDSSF 240
|||||
QY 446 WANHDSKAYVAGGTSMATPIVAGNVAOLREHFVNKRGITTPKPSLLKALJAGAADGIGY 505
|||||
DB 241 WANHDSKAYVAGGTSMATPIVAGNVAOLREHFVNKRGITTPKPSLLKALJAGAADGIGY 300
|||||
QY 506 PNGNGMGRTYTLKSLNVAAYVNESSLSTSOXATYXFTATAGPKLISLWSDAPASTTA 565
|||||
DB 301 PNGNGMGRTYTLKSLNVAAYVNESSLSTSOXATYXFTATAGPKLISLWSDAPASTTA 360
|||||
QY 566 SYTLVNDLDLVITPAKGTXYVGNDFXPPXXXNMGRNNVENVFINKPGSGTYTIEVOAYN 625
|||||
DB 361 SYTLVNDLDLVITPAKGTXYVGNDFXPPXXXNMGRNNVENVFINKPGSGTYTIEVOAYN 420
|||||
QY 626 VPVGPQXFSLAIVN 639
|||||
DB 421 VPVGPQXFSLAIVN 434
|||||
```

RESULT 4

```
Q9AOR1 PRELIMINARY; PRT: 433 AA.
AC Q9AOR1;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Protease (Fragment).
GN PROD.
OS Bacillus sp. SD521.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=133780;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SD521;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships."
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046405; BAB21268.1; -.
DR HSSP; Q45670; IDBI.
DR InterPro; IPR000209; Peptidase-S8.
DR Pfam; PFO0082; Peptidase_S8; 2.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 433 AA; 45576 MW; 9842DF18FE660DDC CRC64;
```

```
Query Match 63.0%; Score 1952.5; DB 2; Length 433;
Best Local Similarity 86.6%; Pred. No. 3.4e-110;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

QY 206 NDVARGIVKADVAQSSYGLYGOGQIVAAVDGLDTGRNDSMEHAFRGKITALVALGRTN 265
|||||
DB 1 NDVARGIVKADVAQSSYGLYGOGQIVAAVDGLDTGRNDSMEHAFRGKITALVALGRTN 60
|||||
QY 266 NANTDNGHGHVAGSVLGNKTKMGAPQANLVFQSIMDSXGGLGGLPSNLTLFESQAXS 325
|||||
DB 61 NANTDNGHGHVAGSVLGN-ALNKGMAPOANLVFQSIMDSXGGLGGLPSNLTLFESQAWN 119
|||||
QY 326 AGARIHTNSGCAVNVGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNPGTISAPGTAKNAI 385
|||||
DB 120 AGARIHTNSGCAVNVGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNPGTISAPGTAKNAI 179
|||||
QY 386 TVGATENLRSPFGSYADNINNHVAQFSSRGPTKDGRIKPDVMAPTXTLLSARSSLAPDSSF 445
|||||
```

```
DB 180 TVGATENLRSPFSLADNPNHIAQFSSRGATRDGRIPDVTAGCTFLLSARSSLAPDSSF 239
|||||
QY 446 WANHDSKAYVAGGTSMATPIVAGNVAOLREHFVNKRGITTPKPSLLKALJAGAADGIGY 505
|||||
DB 240 WANHDSKAYVAGGTSMATPIVAGNVAOLREHFVNKRGITTPKPSLLKALJAGAADGIGY 299
|||||
QY 506 PNGNGMGRTYTLKSLNVAAYVNESSLSTSOXATYXFTATAGPKLISLWSDAPASTTA 565
|||||
DB 301 PNGNGMGRTYTLKSLNVAAYVNESSLSTSOXATYXFTATAGPKLISLWSDAPASTTA 359
|||||
QY 566 SYTLVNDLDLVITPAKGTXYVGNDFXPPXXXNMGRNNVENVFINKPGSGTYTIEVOAYN 625
|||||
DB 361 SYTLVNDLDLVITPAKGTXYVGNDFXPPXXXNMGRNNVENVFINKPGSGTYTIEVOAYN 419
|||||
QY 626 VPVGPQXFSLAIVN 639
|||||
DB 420 VPVGPQXFSLAIVN 433
|||||
```

RESULT 5

```
Q9AOR4 PRELIMINARY; PRT: 433 AA.
AC Q9AOR4;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Protease (Fragment).
GN PROD.
OS Bacillus sp. D6.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=127889;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D6;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships."
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046402; BAB21265.1; -.
DR HSSP; Q45670; IDBI.
DR InterPro; IPR000209; Peptidase-S8.
DR Pfam; PFO0082; Peptidase_S8; 2.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 433 AA; 45636 MW; 52087E0A2516107F CRC64;
```

```
Query Match 62.8%; Score 1948.5; DB 2; Length 433;
Best Local Similarity 86.4%; Pred. No. 6e-110;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

QY 206 NDVARGIVKADVAQSSYGLYGOGQIVAAVDGLDTGRNDSMEHAFRGKITALVALGRTN 265
|||||
DB 1 NDVARGIVKADVAQSSYGLYGOGQIVAAVDGLDTGRNDSMEHAFRGKITALVALGRTN 60
|||||
QY 266 NANTDNGHGHVAGSVLGNKTKMGAPQANLVFQSIMDSXGGLGGLPSNLTLFESQAXS 325
|||||
DB 61 NANTDNGHGHVAGSVLGN-ALNKGMAPOANLVFQSIMDSXGGLGGLPSNLTLFESQAWN 119
|||||
QY 326 AGARIHTNSGCAVNVGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNPGTISAPGTAKNAI 385
|||||
DB 120 AGARIHTNSGCAVNVGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNPGTISAPGTAKNAI 179
|||||
QY 386 TVGATENLRSPFGSYADNINNHVAQFSSRGPTKDGRIKPDVMAPTXTLLSARSSLAPDSSF 445
|||||
DB 180 TVGATENLRSPFSLADNPNHIAQFSSRGATRDGRIPDVTAGCTFLLSARSSLAPDSSF 239
|||||
```

QY	446	WANHSKAYAMGCTSMATPIVAGNVAOLREHFVKNRGITPKPSLLKAALIAAAGXGLCY	505
DB	240	WANYSKAYAMGCTSMATPIVAGNVAOLREHFVKNRGITPKPSLLKAALIAATVGLCY	299
QY	506	PANGNGKGRVTLDRKSLNVAAYVNESSXLSISOKATYXTTATACKPLKISLWSDADASTTA	565
DB	300	PSGDGKGRVTLDRKSLNVAAYVNEALTLTGCAATLSFQIACKPLKISLWTDADASTTA	359
QY	566	SVTLVNDLDLVTAPNGTXYVGNDEXPEPXXXKMDGRNNVENVFIMXPOSGTYTIEOVAYN	625
DB	360	SVTLVNDLDLVTAPNGKYYVGNDESPYDNMDGRNNVENVFIMAPSGTYTIEOVAYN	419
QY	626	VPVGPQXFSLATVY 639	
DB	420	VPSPQRFSLATVH 433	
RESULT 6			
QY	09AOR2	PRELIMINARY; PRT; 433 AA.	
AC	09AOR2		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	Protease (Fragment).		
GN	PROC.		
OS	Bacillus sp. Y.		
OC	Bacteria; Firmicutes; Bacillus/clostridium group; Bacillales;		
OC	Bacillaceae; Bacillus.		
OX	NCBI_Taxid=133779;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Y;		
FX	MEDLINE=2056675; PubMed=11118284;		
RA	Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,		
RA	Horikoshi K.;		
RT	"Novel oxidatively stable subtilisin-like serine proteases from		
RT	alkaliphilic Bacillus spp.: enzymatic properties, sequences, and		
RL	evolutionary relationships."		
RL	Biochem. Biophys. Res. Commun. 279:313-319(2000).		
DR	EMBL; AB046404; BAB21267.1; -.		
DR	HSSP; Q45670; IDBI.		
DR	InterPro: IPR000209; Peptidase_S8.		
DR	Pfam; PF00082; Peptidase_S8; 2.		
DR	PRINTS; PR00723; SUBTILISIN.		
DR	PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.		
DR	PROSITE; PS00138; SUBTILASE_SER; 1.		
FT	NON_TER 1 1		
FT	FT 433 433		
SO	SEQUENCE 433 AA; 45587 MW; BB1291A803C775AE CRC64;		
Query Match 62.6%; Score 1941.5; DB 2; Length 433;			
Best Local Similarity 86.2%; Pred. No. 1.6e-109;			
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;			
QY	206	NDVARGIYKADYAOSSYSLTGGOGIVAYADGLDGRDSSNHEAFPRKITALVALGRTN	265
DB	1	NDVARGIYKADYAOSSYSLTGGOGIVAYADGLDGRDSSNHEAFRGITALVALALGRTN	60
QY	266	NANDTNGHGTIVAGSVLGNXKTKMGAPQANLVEFOSINDSXGGLGGLPSNLDTLFSQAXS	325
DB	61	NASDENGHTIVAGSVLGN-ALNKGMAPQANLVEFOSINDSXGGLGGLPSNLDTLFSQAXN	119
QY	326	AGARHTHTSKGAANVGAATTTDSRNVDDYVRKNDMTITLPAAGNEXNGGTISAPGTAKNAI	385
DB	120	AGARHTHTSKGAPVAVGATYANSROVDEVVRKNDMTITLPAAGNEGNSGTISAPGTAKNAI	179
QY	386	TVGATENTRPSGSGAYADININVAOFSRGPDKGIRKDVNAPGTXIISARSLSLAPDSSE	445
DB	180	TVGATENTRPSFGSLADPNPHIAQFSSSGATRDGRIKEDVTAPEFTIISARSLSLAPDSSE	239
QY	446	WANHSKAYAMGCTSMATPIVAGNVAOLREHFVKNRGITPKPSLLKAALIAAAGXGLCY	505

[illegible]

```

OY 410 FSSNGPTKGRIRPDVMAFGTYILSARSSLA-----PDSEFWANHDSTKAYVNGG 458
DB 508 FSSNGPTKGRIRPDVMAFGTYILSARSSLA-----PDSEFWANHDSTKAYVNGG 650
OY 459 TSMATPIVAGNVAQLREH-----FYXNRGITRPEKPLKALIAAGADUXGLGYP-- 506
DB 651 TSMATPIVAGNVAQLREH-----FYXNRGITRPEKPLKALIAAGADUXGLGYP-- 710
OY 507 -----NGNCGMCHVTLDKSLNVAYNES----- 530
DB 711 STNTNPNNAFEDTFAGANFVQMGSLRMEWL---YVESGKVRKPSRWVGIGELGKDK 767
OY 531 -----XLISTOKATYXFT-----ATAGKP-LKLSLWSDADASTASVTLVND 574
DB 768 ASNMKEVSLSTGONVSYCTFKPSSSGNSGCIPIRYATVLTWDPSPYSGAKLNLVNM 827
OY 575 LVIT-----APNGXYVG-NDFAKXPXAXMMDGRNVEVF---INXQSGITITE 620
DB 828 LTMNTSEPIFYNSGSGSYNGTGTLPLO---DSINVEGITTYPTNKTSEISFRT 884
OY 621 VQAINPVGPOXFS 634
DB 885 IAGTNIPIGPONS 898

RESULT 8
OY 09GTN7 PRELIMINARY; PRT: 1702 AA.
AC 09GTN7;
DR 01-MAR-2001 (TREMBLrel. 16, Created)
DR 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Taga.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA "Taga, a putative serine protease/ABC transporter of Dictyostelium
RT that is expressed at the onset of development and is required for the
RL differentiation of a subpopulation of prespore cells."
DR EMBL; AF263455; AAG11416.1; -.
DR HSSP; P13569; INBD.
DR Interpro; IPR003593; AAA_Atpase.
DR Interpro; IPR001140; ABCtransporter.
DR Interpro; IPR003439; ABC_transporter.
DR Interpro; IPR002029; Peptidase_S8.
DR Pfam; PF00664; ABC_membrane_1.
DR Pfam; PF00005; ABC_tran_1.
DR PRINTS; PR00723; SUPRILISIN.
DR PRODOM; PD000006; ABC_transporter_1.
DR SMART; SM00382; AAA_1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; 1702 AA; 187103 MW; 4A67716303CE7131 CRC64;
SQ
SEQUENCE 1702 AA; 187103 MW; 4A67716303CE7131 CRC64;

Query Match 15.4%; Score 478; DB 5; Length 1702;
Best Local Similarity 21.9%; Pred. No. 1,7e-20;
Matches 167; Conservative 125; Mismatches 261; Indels 210; Gaps 26;
OY 75 KKLTVPANRKLIXOFGNPILETKQXL-----EXTGAKLIDYIPDAYIVVEGCD-- 126
DB 78 KKLTVPANRKLIXOFGNPILETKQXL-----EXTGAKLIDYIPDAYIVVEGCD-- 126
OY 127 -----VXSSXXIXIEHVESVEPYLPYXIDP----- 151
DB 138 NDNNNNKILINRKLKELIPSIQWAKLEPRKLSPLFKONQFQGNQNDQDLKATYYHE 197
OY 152 -----QLTFKASXLYVAXALDITKQXNKE-----VQLR-----GIEIXAQQXXSND 192

```

```

DB 198 NSNOQSDNDNTIISSESLTVEKELISNNNNNNNVLITVNNKSLSESTIYKISTRS 257
OY 193 VXYITAKPEYKYV-----NDVARGIVKADVAQS-----YGLYGQGVAVADPTGLT 240
DB 258 LVYIWEPSSSKLLKTPSKNKAFAHYSIOSGSASTTSPIMDVIGIKDGEIVGACADVTGDI 317
OY 241 GR--NDSSMHEAFRQKITALYALGRTNANADTNGHTHVASGLVNGXTN-----KGM 291
DB 318 NHCFEVDYTNIGSTHKKIIS--YSSGNGDQIDEIDGHQTHVGTIIGSTTVDPSEVSEFG 376
OY 292 APOANLYFOSINDSGXGGLGPSNLQTLFQAXSAGARHTNSMGAANVAGATVDSRVD 351
DB 377 APNSKAFVLDLQVSGNSGLSIOSTNLTAYOSTYDQNAKVCHADAMNSNIGPFTYGTETID 436
OY 352 DYARKN-DMTILEAGNEXPNG-----GISAPGAKNALTGATENLRPSRGSYADIN 406
DB 437 RFQWDHEDFLVYASGNVNVNFGNSITTLSEOSTSKNSLVGSSN--QPS-STYLSIDY 493
OY 407 ----- 406
DB 494 WMDFIYNSIRTSVCTQGSIGITCSDVPTQTSVDIQTCCSNPIIAKICCTEIQOQ 553
OY 407 -----VAOFFSKGPTKGRIRPDVMAFGTYILSARSSLAPOSSFWANH----- 449
DB 554 YQNSTVYSEFTPSLFSGVGPTSDGRKLPDLAAGSPITISR-SLQPSSTI--NHCSPT 610
OY 450 ----DSKYAVMGTSMAPIVAGNVAQLRE-----HFVKNR-----GITPKPSLLKALI 495
DB 611 SGITSAALIMEGSSQAAAVATSAVLRQYRDRGYRINKVNSVGFQPSASLYKALI 670
OY 496 AGAA---DKLQCPNGNOGCVTLTDKSLNVAYNES-----XLISTOKATY 540
DB 671 NTASINVDSTLEY---SQGFNQLSKLITTTNAGTSLDIPSTIEKADPIINGETNSY 727
OY 541 XFRATGKPLKSLVSDAPASTASVTLVNDLIVTA-PNC--TXVYGNQ---FAXPX 594
DB 728 CFSLDKADIDITLVMTDPAGSPSLFTLVNNDLALAVNDELSTYSGNSSETIFKNTS 787
OY 595 XXMMDGRNVEVFINKPQSGTYTLEOVANVPGPOXSLAI 637
DB 788 QVLEFDLNNVEYIRIKADPAGISGVKIEGNIIVIPNOSVSVY 830

RESULT 9
OY 08RBJ2 PRELIMINARY; PRT: 561 AA.
AC 08RBJ2;
DR 01-JUN-2002 (TREMBLrel. 21, Created)
DR 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Subtilisin-like serine proteases.
CN APRE2 OR TTE0824.
OS Thermoaerobacter tengcongensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Thermoaerobacteriales; Thermoaerobacteriaceae; Thermoaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-MB4T / JCM11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome."
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013049; AAM24081.1; -.
KW Protease; Complete proteome.
SQ
SEQUENCE 561 AA; 59968 MW; BA9C5C52F7083A18 CRC64;

Query Match 13.3%; Score 412; DB 16; Length 561;
Best Local Similarity 27.6%; Pred. No. 4,1e-17;
Matches 142; Conservative 74; Mismatches 177; Indels 122; Gaps 19;

```

QY	173	NKEVLQGEIXIAOXXXXNDXYITLAPDEKVMMDVARGVVKDY	217
Db	86	NKPVSADIFITAK-----NIGKNENIKHRKIIPSISIANLTKSOINVLSKLEIYKOIEYD	140
QY	218	-----AOSSYGLYGOGQIVAVADOTGLDTRGSDSMHEAFRGKITALLY	259
Db	141	EPVYATLDATKWEGITKARSDFEVTKRNITIALIDPDIIGNHNHDS-----GSKI----	191
QY	260	ALGFTNNAN-----DTNGCHGTAVGSLGNGXN-----KMAPRANVPOSIMDS--XGC	308
Db	192	-IGMKDFLNKNTPTYYDDHGHTHVAISAAGIAGAGNSYTKGAPAPALLVGIKVLIDANGSGS	250
QY	309	LGGLPSNLQTFESQAXSAGARIHTNSGCAVNGAYTTDSRNVDDYVRKNDMTILFAGNE	368
Db	251	MSYTAGIDMAVQKKDYGIIVILSLGTSTSSDGTSTSLAVRAVADSGIIVYVAAGNS	310
QY	369	XPNNGTISAPCAKNAKNTTGVGTENLRSPESYADNINHAVOFSRSGTCKGRKIPDMAP	428
Db	311	GPAKTYTISPPAAKAKALITVAMAAV-----GELGFLN-----ASFSSRGFTAGGRKIPDIAAP	363
QY	429	GTXILLSARSSILAPDSEFWANHDSKYAAMGTSMTPTIVAGNVALORBEHFKNRNGITPPKS	488
Db	364	GYNITPAK-----ANSVNGEYVYISGSMAKTPPVAGVALMLN-----ANPLNLPNDA	410
QY	489	LKKAALLIAGADXLGLTPNGNGCGRGRTLDKSLNVA-----YVNESSKLST	534
Db	411	--KNIMINSTAASWCPSPKSNVDYGGAGRIDGYEARIVAGNFRGNINIDVPHYYI--SGYLPG	466
QY	535	SOKA-TYXETAT-AGKPLKISLV---MSDAPASTYAGTVLUNDLIVITAPNGTXYYGND	589
Db	467	SRSQTMFWFNNTNNTSYPIATLILIPDMANNP-----DPDIYLPSPSGILIKSS-	515
QY	590	FXFXXXAXXWDSGRNNVENVFINXPOSGTYTTEVQAY	624
Db	516	-----TGTORETTTILPISQGTGYVYKXVSY	541

RESULT	ID	Q80UC9	PRELIMINARY;	PRT;	654	AA.
080UC9	080UC9	080UC9;				
AC	080UC9	080UC9;				
DT	01-JUN-2002	(TREMBLrel, 21, Created)				
DT	01-JUN-2002	(TREMBLrel, 21, Last sequence update)				
DT	01-JUN-2002	(TREMBLrel, 21, Last annotation update)				
DE	Alkaline serine protease.					
GN	PR1670.					
OS	Pyrococcus furiosus.					
OC	Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;					
OC	Pyrococcus.					
OX	NCBI_TaxId=2261;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;					
RA	Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;					
RT	"The complete sequence of the Pyrococcus furiosus genome."					
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.					
DR	EMBL: A010265; AAL81794.1; -					
KW	Protease; Complete proteome.					
SQ	SEQUENCE 654 AA; 70230 MW; 1CB145A5F505DB34 CRC64;					

Query Match	13.2%	Score 408	DB 17	Length 654
Best Local Similarity	25.9%	Pred. No. 8.7e-17		
Matches 162	Conservative 80	Mismatches 210	Indels 174	Gaps 26

```

Qy      62 VESEVNRKXKJLKLKJLEVPANNKLH--XQFNPLIEETKQXJEXGATILYIPDYAV 119
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      35 VEKNVGLTPELFKRIQKLNMEBISVIYVFEENREKEINARVLEMGAV-----RY 87

Qy      120 IVEVEGDVXKXXXIIEHVESVEPLIPXYIXIDPOLFTGASGLYKAXALDTRKQNKVEYOLR 179
      88 V-----YHIIPIAL--AADLKVRDILLVIGLTGGRKLS 118

```

```

0Y 180 G1EX1A0XXXNDXYITAKREYEVAMNDVARG1YKADVAOSSGVLXGQOGIVAVAPDGLD 239
Db 119 GYRF0E0BYK-----YVSALEBGLDEBSAQWATTVMNLGID--GSGITIGIDIDGID 170
0Y 240 TGRDSSMHEAFRGKITALYALGRTNNAN-----DTNGCHTHVAGSVLGNXHTN---K 289
Db 171 -----ASHPDLOGKV-----IGWVDVWNGSKSPYDDHGHGTWAS1A0TGAASNGKYK 219
0Y 290 GMAP0ANLVFQSIM--DSXGGLGGLPSNL0TLEFQAXSAGARIHTNMGW-----337
Db 220 GMAPAKIAGIKVLCADSGSISITTINGEVAMDVNRKRYIKV1NLSLGSOSSDGDAL 279
0Y 338 --AVNGAYTDSRVNDVYVRKNDMTLEPAENGXPNGGTSAPGTAKNAITVGATENLRP 395
Db 280 SQAVNAAMD-----GLVVVVAAGNSGPNKRTTIGSPLAASKVITVGA-----321
0Y 396 SFGSYADININHAOFSSRGPTKKDRIPKDVWAPGTXTLSRSSLIAPDSFWAMHDSKYAY 455
Db 322 -----VDKYDITTFSSNGPFIADRLKPEYVAPBNMTI1AARAS--GTSMG0PINDYTTA 373
0Y 456 MCGTSMATPIVAGNVAOLREHFVKNRGITPK--PSLLKAAL1AGA-----ADXXGLGY 505
Db 374 APGTSMATPIHVAAG1AALLDQ-----AHPSPWPDKVKTALETIADIVKPEDEIAD--1AY 424
0Y 506 PNGNGMGRVYLDKSLINAYVNESSXLSTOSKA-----TYXFATACKPLKISLVMSDAP 560
Db 425 -----GAGRVAAYKAIN--YDNYAKLVFTGIVAKKGSQTHQFV1SGASFTATL1YMNAN 477
0Y 561 ASTPASVYLVNDLDELVIYAPNG-----TXVYGNDEXPXPXXXXNMGDGRNVENVFINXPO 613
Db 478 -----SDLDLYLDPNGNOVDYSYTAYG-----FEKVGYNPT 511
0Y 614 SGTYYIEQAVANVPYQGXESLAVN 639
Db 512 DGTWITKVAYS--GSAN1OVDDV 534

```

```

RESULT 11
O9FBZ4
O9FBZ4 PRELIMINARY; PRT; 1239 AA.
AC O9FBZ4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative secreted peptidase.
GN SC07188 OR SC8A11.16c.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces
OX NCBI_TaxID=1902;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Saunders D.C., Harris D.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

```

RA Kinashih, Hopwood D.A.:
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RN
RP
RC
RA STRAIN-A3(2) / M145,
Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,


```

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrett B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL391041; CAC01588.1; -.
DR HSSP: Q99405; IMPT.
DR InterPro: IPR0003137; PA.
DR InterPro: IPR000209; Peptidase_s8.
DR Pfam: PF02225; PA: 1.
DR Pfam: PF00082; Peptidase_s8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
SQ SEQUENCE 1239 AA; 128505 MW; 8F5E9AC68EB1260A CRC64;

Query Match 12.6%; Score 391; DB 16; Length 1239;
Best Local Similarity 30.9%; Pred. No. 2e-15;
Matches 143; Conservative 51; Mismatches 187; Indels 82; Gaps 16;

OY 213 VKADVAQSSY-----GLYGGQIVAVADTGLDGRNDSMHEAFRGKITALYALG 262
DB 219 VEADLDSTAOIGAPRAWAGGTGQGVAVLDTGVAG-----HPDLADRIAMRQSFV 272
OY 263 RTNNANDTNGHGTAVGSLVNGXTN---KGMAPQANVLFOSIMDSXGGGLPMLQRT 318
DB 273 PDENTDDRCRGHCTHVASTIACTGAASGKKGVAFCARLSIKYVDN--SGRQJSMTLAA 321
OY 319 LFSQAXSAGARLHTNSMGAANCAVYTD--SRNVDYVRNDNTIIFAGNEXPRNGTISA 377
DB 332 MEMAAVERIAKLVNMSLSCGSDSPMSRAVDRLSAOTGALFVYAGN--GGEAGSIGA 390
OY 332 PCTAKNATVGTATENLRPFSGSYADNINMVAOFSKGPDKRIPDVMAAGTXYLSARS 437
DB 391 PGVAISALTVGA-----VDATDLTAPSSSGCPVVDGALKPEITAGVGIILAA-- 437
OY 438 SLAPDSFWAMHDSKYAVMGSTMAPIYVAGNVAQLREHFVNKRGIPKPSILKAAL--- 494
DB 438 ----NSSFAAGNGAYOSLSTGSTMATPHVAGAAALL-----AARPLSGSALKQV 484
OY 495 IGAADXGLGYPNGNGMGRTYLDKSLN-----VAVNNESSXLSTISOKATYFTATAG 547
DB 485 LASSSHRTPRDADFQAGSGRDYDAVRAGVYASATAYAPGSSPGPYRRLVYTNITGAA 544
OY 548 KPLKISLVMSDA-----PASTTASVTLVNDLDTLVITAPNGTYVGNDFEYXP 593
DB 545 VTLELSVATATAPGCVFRLSASRYVPAAGTADVTLLTIDGS---GSAAGRAYSGQILLAD 601
OY 594 XXXXNDGRRNENVFINKPQSGTYTTEVQ--AYNVEVGFQXFSL 635
DB 602 A----DARNVAHTAVSAGPVRAKRLTVHFKDADGNPV--PGVFDL 639

RESULT 12
O9PFC06 PRELIMINARY; PRT: 1253 AA.
AC O9PFC06
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE Putative 2002 (Tremblrel. 21, Last annotation update)
GN SC07176 OS: SC8B11.04C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteriales; Actinobacteria; Actinobacteridae;
OC Actinomycetiales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;

[1]
SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Saunders D.C., Harris D.;
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RC Cerdeno A.M., Parkhill J., Barrett B.G., Rajandream M.A.;
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE-97000351; PubMed-8843436;
RA Kadenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
RA Knaashl H., Hopwood D.A.;
RT A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RP M01. Microbiol. 21:77-96(1996).
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RA Penley S.D., Chatter K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrett B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL391041; CAC01576.1; -.
DR HSSP: Q99405; IMPT.
DR InterPro: IPR0003137; PA.
DR InterPro: IPR000209; Peptidase_s8.
DR Pfam: PF02225; PA: 1.
DR Pfam: PF00082; Peptidase_s8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
SQ SEQUENCE 1253 AA; 130971 MW; AA69B417EEEDB89 CRC64;

Query Match 11.9%; Score 368.5; DB 16; Length 1253;
Best Local Similarity 28.0%; Pred. No. 4.7e-14;
Matches 142; Conservative 55; Mismatches 169; Indels 141; Gaps 20;

OY 205 NNDVARGIVKADVAQSSYGLYGGQIVAVADTGLDGRNDSMHEAFRGKITALYALGRT 264
DB 219 LSDTYTAQIGADV--WGGNTGEGVAVLDTGVAG-----HPDRGRLATATASFVDP 270
OY 265 NNANDTNGHGTAVGSLVNGXTN---KGMAPQANVLFOSIMDSXG----- 307
DB 271 QDVYDRNCHGTHVASTIAGTGAASGVEKGVAFGASLHTGVLNDSGSDSWVLAMGEW 330
OY 308 -----GLGGIPSMQLTFLFSQAXSAGARLHTNSMGAANCAVYTTDSRNVDDYR 355
DB 331 AYRDGHAKIVMSLGDSPITDGTDPLEAV-----NMLSAETGA----- 368
OY 366 KNDWTILFAAGNENXPGTISAPGTAKNATVGAENLRPFSGYADNINMVAOFSRRP 415
DB 369 ----LFLVAAGNSGPFAVYVYGTPTAADAALTVGAVNG--PQKQ-----VDLADPSSKRP 417
OY 416 -TKDGRTRPDVMAPGXYIISARSILAPDSFWAMHDSKYAVMGSTMAPIYVAGNVAQLR 474
DB 418 RYGDNAVKPDLTAPGCVGLAARSRYAPES-----EGAYOSLSTGSTMATPHVAGAAALLA 471
OY 475 EHFVKNKGITPRKPSLKAALIGAADXGLGYPNGNGMGRTYLDKSLNVAVYVNESSXLST 534
DB 472 AENPRDTG-----QRLKELVGTGTAGTGRFSP--FDAGSGNV-----DVAAVASTLLAS 519

```

```

QY 535 SQ--KAY-----XFPATGKPLKLTWSDA-----PAST 563
Db 520 GDAFAQAHYPTPGQTVRBDVTFYNSGAPVALDLALSPALPECLFTLEAQVTPAHG 579
QY 564 TASTYLVNDL-----LVITAPNGXHYVGNDFXXPKXNMUGRNVEVFL--- 609
Db 560 TASTGVYIHLDAEDNGCATRLVLAAGDQAVLA---RTTPGVNKEGRRAITALTAKDH 635
QY 610 -NKFQSGTYTLEVOAIVNVPVGPQXSL 635
Db 636 HDKPLSTGVTLKDVERN--TAPKIVSY 660

```

RESULT 13			
P95684	PRELIMINARY:	PRT:	1102 AA.
ID	P95684		
AC	P95684		
DT	01-MAY-1997 (TREMBLrel. 03. Created)		
DT	01-MAY-1997 (TREMBLrel. 03. Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21. Last annotation update)		
DE	Subtilisin-like protease.		
OS	Streptomyces albobacillus.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
OC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces		
OX	NOBL_TaxID=1887;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-S-3253;		
RX	MEDLINE=97144528; PubMed=8990295;		
RA	Suzuki M., Tachuchi S., Yamada S., Kojima S., Miura K., Momose H.,		
RT	"A novel member of the subtilisin-like protease family from		
RT	Streptomyces albobacillus."		
RL	J. Bacteriol. 179:430-438(1997).		
DR	EMBL: D83672; BAA12040.1; ..		
DR	HSSP: P00782; 2SER.		
DR	InterPro: IPR002860; GH_BNR.		
DR	InterPro: IPR000209; Peptidase_S8.		
DR	Pfam: PF02012; BNR; 2.		
DR	Pfam: PFO0082; Peptidase_S8; 1.		
DR	PRINTS: PR00723; SUBTILISIN.		
DR	PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.		
DR	PROSITE: PS00137; SUBTILASE_HIS; 1.		
DR	PROSITE: PS00138; SUBTILASE_SER; 1.		
DR	Protease.		
SO	SEQUENCE	1102 AA; 114128 MW; F9E4AD2590FE559E CRC64;	

```

Db 453 ---MKVAELKGLTASTKRG--KYMPEEGSGRVQVDATITQVTAAEPVLSFVGQM 503
Qy 534 -----TSOKATXFPTAACKPLKISLVWSD-----APAS--TTASVATVINDLUITA 579
Db 506 PHADDKPVTKLITRLNLGTEVYTLKLTSTAGPGRKAAPAGFLGLSTL-----TY 557
Qy 560 P-NGTYXYGNDPFXPFXKXNMDDRRNVENVFINXQS-----GTYITEVQAYNV 626
Db 558 PANGTAASDVYDATRLGAGVADGTSAAVYVATGAGQSVTAANAVEREESTNV 609

```

ID	09KBJ7	PRELIMINARY;	PRT: 444 AA.
AC	09KBJ7;		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Intracellular alkaline serine protease.		
GN	APRX OR BH1930.		
OS	Bacillus halodurans.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;		
OC	Bacillaceae; Bacillus.		
OX	NCBI_TaxID=86665;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-C-125 / JCM 9153;		
RX	MEDLINE=20512582; PubMed=11058132;		
RA	Takam H., Nakasone K., Takaki Y., Maeno G., Saseki R., Masui N.,		
RA	Fuji F., Hirata C., Nakamura Y., Ogasawara N., Kuhara S.,		
RA	Horikoshi K.;		
RT	"Complete genome sequence of the alkaliphilic bacterium Bacillus		
RT	halodurans and genomic sequence comparison with Bacillus subtilis.";		
RL	Nucleic Acids Res. 28:4317-4331(2000).		
DR	EMBL: AP001513; BAB05649.1; -		
DR	HSSP: A99405; 1MPM		
DR	InterPro: IPR001128; Cytochrome_P450.		
DR	InterPro: IPR002029; Peptidase_S8.		
DR	Pfam: PF00082; Peptidase_S8; 1.		
DR	PRINTS: PR00723; SUBTILISIN.		
DR	PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.		
DR	PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.		
DR	PROSITE: PS00137; SUBTILASE_HTS; 1.		
DR	PROSITE: PS00138; SUBTILASE_SRR; 1.		
DR	Protease; Complete proteome.		
QC	SEQUENCE 444 AA; 48916 MW; 3805F81053A21F2F CRC64;		

```

Query Match          11.5%; Score 356; DB 2: Length 1102;
Best Local Similarity 30.5%; Pred. No. 2.3e-13;
Matches 144; Conservative 55; Mismatches 185; Indels 88; Gaps 19.

QY 203 KVMNDVAR----GIYKADVAQS-----SYLHGCGQIVAAVADTLDTGRNDSSMH 248
      :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :|
Db 178 RTASVAVARVMDGVKAKSLDTSVSGIGTPKMEAGYDGGVAVIAVLDTVD-----AHH 231
      :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :|

QY 249 EAFRGKITALYALGHTNNANDPTNGHGTTHVAGSVLNGX---TNGMAPQANLVFQSID 304
      :  :| | :  :| | :  :| | | | | | | | | | | | | | | | | | | | |
Db 232 PDLKQGVATSKNFTSAPITPTGGVYGHGTTHVASTAAGTGAGSKCTYKGVANAGATILNKQVLD 291
      :  :| | :  :| | :  :| | | | | | | | | | | | | | | | | | | | |

QY 305 SXGGLGGLPSMLQTLFSQAXSAGARIHTNSWGAAVNGAYTTDSRVNDYVVR--NDMTIL 362
      :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :|
Db 292 DAG--FGDDSGILGMEWMAAAGGADIVNMSLG---GMDTPPTDPLEAVDLSAKETLL 345
      :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :|

QY 363 F--AAGNEXPMNGGTSAPGTAKNAITVGATENTLRFSPGSYADINHHVAQFSSGRG--TKDG 419
      :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :|
Db 346 FAIAGNNGEPQ--STGSPGSAADSALTVA-----YDDKDKLADFSSSTGPRLDGG 392
      :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :|

QY 420 RIKPRVMAPEGTYILSARSLADPSSFMANHDSKYIYMGTSNATPIPVAGNVAQLREHYVK 479
      :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :|
Db 393 AVKPDLTAPGVDTFAASKGNDIAKEVEKEKPGAGYMTISGTSMAATPHVAGAAALLKQOHPD 452
      :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :|

QY 480 NRGITPKRSILKAALIAQAADYGLGTPMGNGMGWGRVTLDDKSLINVAIVNNESSKLS----- 533
      :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :| | :  :|

```

```

Query Match 11.3%; Score 350.5; DB 16; Length 444;
Best Local Similarity 30.2%; Pred. No. 1.6e-13;
Matches 114; Conservative 54; Mismatches 127; Indels 83; Gaps 17;

QY 182 EXIAQXXXXXNDPVXYITAKPEYKVMNDVARGIVKA-DVAOSSGLYGOCQIVAVDTGLDT 240
Db 100 ESLQEMLYCKDIDRKITYLAREVHALDPTVESQAQAEVIRNGETTLGKDYITAVIDGT-- 157
QY 241 GRNDSMEHEARGKITALLY-ALGRTNNANDTNGSHTHVAGSVLGCXTN----KGMAPQ 295
Db 158 -----YPHEDELRKAKAVDEYNQAEPEEYDDNGHGTHCAGDAAGANGASSDQYRGCPAPBA 212
QY 296 NLVFOSSINDSXGGLGGLPSNLQTLFSSQAXSAGARITHNTSMGAAVNGAAYTTDSRNV----- 350
Db 213 NVIGIKVYLKQ-QMGSLTESIMQV-----EMCQYNEHRDDPHIHLISML 257
QY 351 -----DDYVR-----KNDMTILFAGNEXPNNGTISAPGTAKNAITVGATEN 392
Db 258 GGQALPYENEGEDDPVNRILVEEAMNNGITIVCYAAGNSGFDQAQIASPGVSEKIVITGALDD 317
QY 393 LRPFSGTADNINHYAOFSSRGPTKDGRIKDPVMAPTGYILISARSLAPDSSF-----WA 447
Db 318 -RDTTDRRDD--VAPFSSRGPTTYGKRPKPDILAPGVNISLRS--PNSFYDKIKQGS 370
QY 448 NHDSKYAMGTSNATPYAVNGVAVOLREHFYKNGKITRPSLKLKALLAGA---AD----- 500

```

```

DB 371 RVGSHYTMSTGSMATPACVAVLALMLOH---EPNLTPDE--VKTRLMGSTRMADRDPN 425
QY 501 -XGLGY-----PNGNO 510
DB 426 VYGAGYISARCAIIPNSE 443

RESULT 15
O9RL54
ID O9RL54 PRELIMINARY: PRT: 1245 AA.
AC O9RL54;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE Probable secreted peptidase.
GN SC00432 OR SCF51A.10.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Murphy L., Harris D.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RP [2]
RC STRAIN-A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinasht H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RN [4]
MO1. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RL Nature 417:141-147(2002).
DR EMBL: AL121596; CAB56662.1; -.
DR HSSP: P00782; 1SUP.
DR InterPro: IPR003137; PA.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00225; PA.1.
DR Pfam: PF00062; Peptidase_S8.1.
DR PRINTS: PR00723; SUBTILISIN_HIS.
DR PROSITE: PS00137; SUBTILISIN_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILISIN_HIS; 1.
SQ SEQUENCE 1245 AA; 130895 MW; 74EE92DB9CA1DE60 CRC64;

Query Match 10.9%; Score 337; DB 16; Length 1245;
Best Local Similarity 31.5%; Pred. No. 3,7e-12;
Matches 123; Conservative 40; Mismatches 158; Indels 70; Gaps 14;

QY 221 SYGLYGQGIIVAAVDLTGTGRNDSMHEAFRGKITALYALGRTNANNDPTNGHGHVA-- 278
DB 227 SAGRYGAGVAVVAVLDTGAD-----QSHPDLAGRVAAAKDFSSSGTNDVFGHGHVASI 280

```

```

QY 279 --GSVLGNGXTNKMAPQANLVFOSIMDSXGLGLGPSNLQTLFSSQAXSAGARIHTNSWG 336
DB 281 VGGSGAASGGSGRQVAPARLLVKGVLGDDG--FGSESGIAGMEMAAGADAVVNNSLG 338
QY 337 A--AVNGCAVTTD--SRNVDYVRKKNDMTILFAAGNEXPNG-GTISAPGTAKNAITVGATE 391
DB 339 SSGATDG---TDPMSQALNDLSRRTGTLFVVAACNCEGCPPTVSGPAAADALTVGA-- 393
QY 392 NLRPSEGSYADNINHYAOFSSRGP-TKDRIRKPDYMAPGTXLISARSLAPSPFWANHD 450
DB 394 -----VDRDLSLAPSSSRGPRLGDDAVKPDVTAAPGVYIAR---AAGSAMGDPVD 441
QY 451 SKYVMGCTSMATPIYAGVAVQALREHFVKNRGITPKPSLLKALLINGAADXGLGTPNGNO 510
DB 442 EHYTAASGTSMATPVAAGAAULLAQRHPDMTGAOLKDALISTAVTVD-----GOKVTEQ 495
QY 511 GWGRV-----TLDKSLNAVAVNESSXLSSTOKATVYKPTAT 545
DB 496 GGGRIDVRAAGLGAVTATGTLVMGPFITSRDTEPVTISRVRTNSDDEDTLLSLAVELATEG 555
QY 546 AGKPLKISL-VMSDA---PASTTASVTLVND 572
DB 556 GKAPAECSARLGSDSVRVPAGSSAEVPLTVD 586

```

Search completed: April 1, 2003, 12:40:54
Job time : 64.4527 secs

THIS PAGE BLANK (uspto)

THIS PAGE BLANK (uspto)